Matching DSGE Models To Data With Applications To Fiscal And Robust Monetary Policy

Three Essays In Dynamic Macroeconomics

DISSERTATION

zur Erlangung des akademischen Grades
doctor rerum politicarum
(Dr. rer. pol.)

eingereicht an der
Wirtschaftswissenschaftlichen Fakultät
Humboldt-Universität zu Berlin

von
Herr Alexander Kriwoluzky, (M.A.)
geboren am 6.7.1978 in Berlin

Präsident der Humboldt-Universität zu Berlin:
Prof. Dr. Dr. h.c. Christoph Markschies

Dekan der Wirtschaftswissenschaftlichen Fakultät:
Prof. Oliver Günther, Ph.D.

Gutachter:

1. Prof. Harald Uhlig, Ph.D.
2. Prof. Bartosz Mackowiak, Ph.D.

Tag des Kolloquiums: 9. Juli 2009

Abstract

This thesis is concerned with three questions: first, how can the effects macroeconomic policy has on the economy in general be estimated? Second, what are the effects of a pre-announced increase in government expenditures? Third, how should monetary policy be conducted, if the policymaker faces uncertainty about the economic environment.

In the first chapter I suggest to estimate the effects of an exogenous disturbance on the economy by considering the parameter distributions of a Vector Autoregression (VAR) model and a Dynamic Stochastic General Equilibrium (DSGE) model jointly. This allows to resolve the major issue a researcher has to deal with when working with a VAR model and a DSGE model: the identification of the VAR model and the potential misspecification of the DSGE model.

The second chapter applies the methodology presented in the preceding chapter to investigate the effects of a pre-announced change in government expenditure on private consumption and real wages. The shock is identified by exploiting its pre-announced nature, i.e. different signs of the responses in endogenous variables during the announcement and after the realization of the shock. Private consumption is found to respond negatively during the announcement period and positively after the realization. The reaction of real wages is positive on impact and positive for two quarters after the realization.

In the last chapter 'Optimal Policy Under Model Uncertainty: A Structural-Bayesian Estimation Approach' I investigate jointly with Christian Stoltenberg how policy should optimally be conducted when the policymaker is faced with uncertainty about the economic environment. The standard procedure is to specify a prior over the parameter space ignoring the status of some sub-models. We propose a procedure that ensures that the specified set of sub-models is not discarded too easily. We find that optimal policy based on our procedure leads to welfare gains compared to the standard practice.

Keywords:
Bayesian Model Estimation, Vector Autoregression, Identification, Government expenditure shock, Optimal monetary policy, Model Uncertainty, Non-invertibility

Zusammenfassung

Diese Doktorarbeit untersucht drei Fragestellungen. Erstens, wie die Wirkung von plötzlichen Änderungen exogener Faktoren auf endogene Variablen empirisch im Allgemeinen zu bestimmen ist. Zweitens, welche Effekte eine Erhöhung der Staatsausgaben im Speziellen hat. Drittens, wie optimale Geldpolitik bestimmt werden kann, wenn der Entscheider keine eindeutigen Modelle für die ökonomischen Rahmenbedingungen hat.

Im ersten Kapitel entwickele ich eine Methode, mithilfe derer die Effekte von plötzlichen Änderungen exogener Faktoren auf endogene Variablen geschätzt werden können. Dazu wird die gemeinsame Verteilung von Parametern einer Vektor Autoregression (VAR) und eines stochastischen allgemeinen Gleichgewichtsmodelles (DSGE) bestimmt. Auf diese Weise können zentrale Probleme gelöst werden: das Identifikationsproblem der VAR und eine mögliche Misspezifikation des DSGE Modells.

Im zweitem Kapitel wende ich die Methode aus dem ersten Kapitel an, um den Effekt einer angekündigten Erhöhung der Staatsausgaben auf den privaten Konsum und die Reallöhne zu untersuchen. Die Identifikation beruht auf der Einsicht, dass endogene Variablen, oft qualitative Unterschiede in der Periode der Ankündigung und nach der Realisation zeigen. Die Ergebnisse zeigen, dass der private Konsum negativ im Zeitraum der Ankündigung reagiert und positiv nach der Realisation. Reallöhne steigen zum Zeitpunkt der Ankündigung und sind positiv für zwei Perioden nach der Realisation.

Im abschließendem Kapitel untersuche ich gemeinsam mit Christian Stoltenberg, wie Geldpolitik gesteuert werden sollte, wenn die Modellierung der Ökonomie unsicher ist. Wenn ein Modell um einen Parameter erweitert wird, kann das Modell dadurch so verändert werden, dass sich die Politikempfehlungen zwischen dem ursprünglichen und dem neuen Modell unterscheiden. Oft wird aber lediglich das erweiterte Modell betrachtet. Wir schlagen eine Methode vor, die beiden Modellen Rechnung trägt und somit zu einer besseren Politik führt.

Schlagwörter:
Bayesianische Modellschätzung, Vektor Autoregression, Identifizierung,
Staatsausgabenerhöhung, Optimale Geldpolitik, Modellunsicherheit,
Nicht-invertibilität

Acknowledgements

This thesis is the result of my work over the past four years. During this time, many people supported me and contributed to it. Here, I want to mention them one by one.

Most of all, I am indebted to my thesis supervisor Harald Uhlig. It is due to his inspiring courses on dynamic macroeconomics that I have developed such an interest in this field. Later, I was lucky enough to have the opportunity to become his Ph.D. student. His numerous suggestions and comments have been of enormous influence to my work. I am also very grateful to him for introducing me to Chris Sims and initiating a stay at Princeton University. Besides employing me at collaborative research center 649, he supported my applications to external funding, for which I am also thankful.

Furthermore, I want to thank Bartosz Mackowiak, my second supervisor, for taking his time and discussing some issues of my work in depth. I have learned a lot talking to him and working on a joint project which unfortunately did not make it into the thesis, but I am confident that we will finish it in the near future.

On Harald Uhlig's initiative I was invited by Chris Sims to visit Princeton University. This thesis benefited substantially from his comments, chapter 2 having been mostly worked out during my stay at Princeton University. I also had the opportunity to talk to Noah Williams about chapter 4 of the thesis, which is closely related to his work. So I owe him a lot, too.

I would further like to stress the inspiring and fruitful atmosphere of the group of other students I have been working with during the past years. Especially, I want to thank Christian Stoltenberg, who is also co-author of chapter 4, Martin Kliem, Holger Gerhardt, Stefan Ried and Mathias Trabandt. Fortunately, we did not only work. Thanks for that, too. I am also grateful to Susann Roethke for her administrative support and to Jan Auerbach, Patrick Habscheid and Simon Roesel for their research assistance.

For funding I want to particularly thank the collaborative research center 649 at Humboldt University Berlin. Moreover, I received funding from the DEKA Bank for two years and a research grant from the German Exchange Service (DAAD) while I was visiting Princeton University.

Besides an academic side of life I am fortunate enough to have a circle of family and friends - a never ending source of support. Above all, I want to thank Silke for being by my side and encouraging me in my work.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Scope of the study

Macroeconomic policy is one of the most important determinants for the well being of a society. Recent examples show how good macroeconomic policy can improve the overall economic conditions: Norway raised the living standard of its inhabitants by carefully managing the wealth after the exploration of vast oil reserves instead of spending it right away. In Western Germany after World War II, the introduction of a stable currency combined with market reforms substantially helped to boost the economy. On the other hand, bad macroeconomic policy can have severe consequences for the well being and stability of a society as the crisis in Argentine showed in the late 1990s. When the Argentine government spent a large amount of the foreign reserves, the resulting devaluation of the Argentine peso and the lock of many savings accounts led to sharp increase of people living in poverty.

Given the impact macroeconomic policy has on the development of today's societies, it is the primary task of macroeconomists to equip policy makers with good advise. It is therefore important to understand the effects of macroeconomic policy and to determine how it should be set optimally. This thesis aims at contributing to these issues and improve the understanding of three areas in macroeconomics: first how the effects of policy innovations can be estimated, second what the effects of an innovation in government spending are and third how monetary policy should be conducted if the policy maker faces uncertainty about the economic environment.

Knowing the precise effects of a policy instrument is of high value for the policy maker. Unfortunately, the estimation of econometric time series models, the vehicle best suited to examine these effects, exhibits a fundamental challenge: Vector Autoregression (VAR) models can only be estimated in a reduced form, where the error terms are correlated and their effects cannot be interpreted in an economically reasonable way. This is only possible in a structural model. However, an estimated reduced form model cannot be uniquely transferred into a corresponding structural model without imposing additional assumptions. The second chapter of the thesis proposes a methodology to identify the structural model by deriving additional restrictions from a dynamic stochastic general equilibrium (DSGE) model.

These restrictions are determined as follows: After an innovation the impulse response functions of the DSGE model are computed and the sign for some variables (but not for the variable of interest) are imposed onto the VAR model, i.e. the structural shock of the VAR model has to satisfy those restrictions. One obstacle is that different parameterizations of a DSGE model yield different sign restrictions. I therefore aim at finding the parameter values fitting the data best. Since the exercise is concerned with the estimation of dynamic effects, the parameter values of the DSGE model are estimated by matching the impulse response functions of the VAR model.

The third chapter puts the methodology at work: I analyze an innovation in government expenditures. The effect of a government expenditure shock is not only intensively debated among politicians each time the economy is confronted with a recession, it also divides macroeconomists on the question whether private consumption and the real wage rise or fall after an increase in government expenditures. In the former case an expansive fiscal policy would support the economic development and would thus be a good policy prescription, in the latter it would not.

Findings in empirical work differ depending on the empirical strategy. I advocate to estimate the effects of a government expenditure shock by taking one crucial aspect of its nature into account: Changes in government expenditures are often pre-announced. Certain variables then might react qualitatively different during the pre-announcement period and after the realization. I employ a DSGE model which is rich enough to account for different effects of an innovation in government expenditures and derive sign restrictions from it to identify the VAR model. The innovation is assumed to be announced two quarters in advance. The results indicate that a positive shock to government expenditures leads to an increase in private consumption and real wages.

The fourth chapter (written jointly with Christian Stoltenberg) does not deal with the effects policy has, but with the question how policy should be conducted optimally if the policy maker is uncertain about the true economic environment. We take the perspective of the central bank and aim at finding a policy rule, i.e. a rule that determines the interest rate depending on inflation and the difference between the actual and the natural rate of output. This rule is optimal if it maximizes welfare of the economy, that is the households' utility. We define two sources of uncertainty: uncertainty about the parameters of the model and uncertainty about the specification of the model. Each combination of parameters and each specification of the model has certain probabilities. We estimate the probabilities and compute a policy rule which maximizes welfare across the specifications weighted by the corresponding probabilities.

We find that a policy maker with a concern for robustness should not simply include all possible features into one model and use this exclusively to find the optimal policy rule. An optimal policy rule derived from this model does no guard against model uncertainty. It still performs poorly across a set of models including smaller versions of it. The policy implications derived from these smaller versions of the model can differ substantially from the implications of the larger model. We show that this needs to be taken this into account by computing a rule optimal across the model space instead of optimal in the largest model. While we stress that specification uncertainty is important to consider, we also find that parameter uncertainty does not have serious consequences: A policy rule determined at the mean of the parameter distribution performs as well as a rule capturing the uncertainty.

To the three topics of this thesis, the estimation of the dynamic effects of innovations in general, the special case of the effects of an innovation in government expenditures and the optimal conduct of monetary policy under model uncertainty, several contributions have already been made. In the following section I will provide a brief overview over the existing literature and juxtapose my work to it.

1.2 Literature review

In this section I am going to review first the literature on the identification of structural VAR models, the estimation of DSGE models and frameworks aiming at the estimation of both jointly in general. Afterwards I will survey existing work on the specific case of the effect of an innovation in government expenditures. Finally, the literature on optimal monetary policy under model uncertainty is presented.

1.2.1 Identification of a structural VAR model

When Sims (1980) introduced VAR models as a tool into macroeconomics he recommended to identify an estimated reduced form VAR model under a recursiveness assumption that imposes a certain order to the variables of the model. This assumption induces that variables ordered first do not respond immediately to innovations in variables ordered later. It is therefore an assumption concerning the timing of the responses. This procedure results in a sufficient number of additional zero restrictions to identify the VAR model. This approach is widespread and used for example in Christiano, Eichenbaum, and Evans (1999) to identify an innovation in monetary policy.

A related identification scheme called contemporaneous restriction was introduced by Sims and Zha (2006). The authors extend the concept of a recursive ordering. The restrictions can be chosen more freely by determining which variables respond contemporaneously to an innovation in another variable and which variables are predetermined.

Blanchard and Quah (1989) suggest to identify a VAR model by distinguishing transitory shocks, i.e. shocks which have no permanent effects on the variables of the VAR, and permanent ones. Since the behavior of the variables is restricted in the long run, this kind of identification is called long run restriction. It is applied for example when studying the effects of a technology shock, arguing that a technology shock is the only shock which has a permanent impact (Francis and Ramey (2005) and Galí (1999)).

In contrast to the identification of a shock by his long run properties Uhlig (2005a), Faust (1998), Dwyer (1998), Canova and Nicolo (2002) and Canova and Pina (1999) suggested to investigate the effects of an innovation by restricting the impulse response functions of the variables directly. Canova and Nicolo (2002) and Canova and Pina (1999) identify a monetary policy shock by restricting the sign of cross correlations, Dwyer (1998) the shape of the response of some variables, and Uhlig (2005a) and Faust (1998) the sign of the response. While Faust (1998) only consider the first period, Uhlig (2005a) restricts the signs of the responses for a longer horizon. In those studies only one innovation is analyzed. Mountford and Uhlig (2005) apply the concept to a combination of shocks. Kociecki (2005) provides a general framework to formulate a distribution over impulse response functions and thus formulate beliefs over their sign and shape.

While the structural VAR model literature employs no information about the date of innovations in monetary and fiscal policy, Romer and Romer (1990) suggest to study the history of decisions of the central bank, i.e. the precise historic record, and use those to study the effects of policy innovations. It is also employed in studies concerning fiscal policy (Ramey and Shapiro (1998) and Edelberg, Eichenbaum, and Fisher (1999)).

All these identification schemes depend on the availability of some common a priori knowledge. In chapter 2 I suggest to consider a DSGE model as the source of this common a priori knowledge and to derive restrictions from it. Since the restrictions depend on the parametrization of the DSGE model, I now review the literature on how to choose the parameters of the model.

1.2.2 Estimation of DSGE models

When Kydland and Prescott (1982) suggested to base macroeconomic analysis on DSGE models they argued that the DSGE model represents only a small fraction of the economy and is not meant to be estimated - in contrast to the large scale models applied in those days. Instead they suggest to calibrate the DSGE model by choosing reasonable values for its parameters found either in microeconomic studies or by fitting the steady state of the DSGE model to long run characteristics in the data. In succeeding papers (Kydland and Prescott (1991) and Kydland (1992)) give a detailed illustration of how the parameter values should be determined.

Their approach came under attack by Hansen and Heckman (1996). They point out that values found in microeconomic studies should not be used for two reasons: either the existing microeconomic studies are to sparse or there exists a wide range of estimates, which leads to a selection bias and therefore to inconsistencies between studies employing a DSGE model as a tool. As a response to the criticism and in order to be able to quantify the uncertainty about the parameter choice, consequent studies estimated the parameters within the DSGE model framework. Thereby two strands of the literature emerged, which differ considerably in their econometric interpretation of the DSGE model. One strand interprets the DSGE model in the spirit of Kydland and Prescott and seeks to match some selected moments as closely as possible. The other strand considers the DSGE model as a full characterization of the observed time series. In the following paragraphs I am going to review both strands more specifically.

Early attempts to estimate DSGE models were made by Lee and Ingram (1991), Canova (1994), Canova (1995) and Christiano and Eichenbaum (1992). The former authors estimated the parameters of the DSGE model by simulated methods of moments (SMM), the latter by generalized methods of moments (GMM). Both approaches build on the original idea by Kydland and Prescott that parameters of the DSGE model can be found by matching moment characteristics in the data. The parameters of the DSGE model are found by minimizing the distance between the moments implied by the DSGE model and those of the data, an estimation procedure introduced in a different context by Hansen (1982).

The idea of matching certain moments of the data was further used in the studies of Rotemberg and Woodford (1998), Christiano, Eichenbaum, and Evans (2005), Ravn, Schmitt-Grohé, and Uribe (2007) and Mertens and Ravn (2008). The authors aim at estimating the parameters of the DSGE model by minimizing the distance between the implied impulse response function of the DSGE model and the impulse response of the VAR model. Applying this estimation strategy, one issue becomes crucial: the identification of the impulse response functions of the VAR model.

While those estimation procedures rely on weak assumptions concerning the ability of the DSGE model to be a representation of the data generating process, other approaches aimed at employing a DSGE model as a full characterization of the observed time series and not only of some moments. This assumption allows to estimate the DSGE model by maximum likelihood. One of the main obstacles to this estimation strategy is that the DSGE model typically is not fully stochastically specified, since the number of structural shocks is smaller than the number of observed variables to be explained. Following Sargent (1989), numerous authors (McGrattan (1994), Hall (1996), Altug (1989), McGrattan, Rogerson, and Wright (1997) and Ireland (2004)) therefore add error terms to the structural equations of the DSGE model. Those error terms, called measurement error terms, exhibit the difficulty that they are hard to interpret economically.

Other studies abstained from including measurement errors and fitted the DSGE model to a small number of time series (DeJong, Ingram, and Whiteman (2000), Kim (2000) and Ireland (2001)). Smets and Wouters (2003) extend a DSGE model with several structural shocks and additional nominal frictions and features in order to confront the DSGE model with seven key macroeconomic time series. Additionally, building upon DeJong et al. (2000), they employ a Bayesian approach, i.e. they combine a prior distribution for the structural parameters with the likelihood and approximate the posterior distribution. Their paper, together with the availability of "straightforward-to-use" computer programs (Dynare) and of the computational power needed to use these programs, paved the way for the estimation of DSGE models using this Bayesian DSGE model estimation methodology.

In this thesis I comply with the weak econometric interpretation of the DSGE model and estimate the parameters of the DSGE model by matching the corresponding impulse response functions of the VAR model. At the same time, the VAR model is estimated using restrictions from the DSGE model, i.e. both models are considered jointly. The next section reviews the related literature.

1.2.3 The DSGE model and the VAR model considered jointly

Doan, Litterman, and Sims (1984) showed that the ability to forecast from a VAR model can be improved by postulating a prior distribution for the parameters. In their work they suggested a random walk. Ingram and Whiteman (1994) advanced this idea by advocating that the prior distribution of the parameters should not be centered at a random walk, but at moments (dummy observations) computed from a DSGE model. DelNegro and Schorfheide (2004) build on this approach. They provide additional insight on how the posterior distribution of the parameters of VAR model can be used to infer on the parameters of the DSGE model.

Furthermore, they identify the VAR model using information from the DSGE model.[1] They employ the rotation matrix of the DSGE model to identify the VAR model. To do so, the DSGE model has to be fully stochastically specified. Furthermore, while one can control for the prior weight of the dummy observations, one cannot control for the prior weight of the implied dynamics of the DSGE. Sims (2006b) extends the idea to augment the VAR model with dummy observations in a more general framework. In his approach, the tightness of the prior can be varied across frequencies and the number of structural shocks does not need to equal the number of observations.

The methodology proposed in this thesis differs from DelNegro and Schorfheide (2004) and Altig et al. (2002) by not using the implied rotation matrix of the DSGE model to identify the VAR model and therefore not requiring the DSGE model to be fully stochastically specified. In contrast to Sims (2006b), I employ the implied sign and shape restrictions (as described in Uhlig (2005a)).

While the preceding sections were concerned with VAR models and how to identify their structural form, the following section will deal with the application of this methodology and its results for a specific innovation: An innovation in government expenditures.

1.2.4 Government expenditure shock

In empirical studies, findings on the effects of a government expenditure shock are twofold depending on the identification scheme employed.

Ramey and Shapiro (1998) use a narrative approach to identify the VAR model. They interpret times of large military buildups in the US, the Korean war, the Vietnam war and the Carter-Reagan buildup, as sudden and unforseen increases in government

[1] Altig, Christiano, Eichenbaum, and Linde (2002) also pursue this road.

expenditures. The resulting reactions of macroeconomic variables to those events are thus interpreted as deviations from normal behavior. They find that output and hours rise, while consumption and real wages fall. Burnside, Eichenbaum, and Fisher (2004) employ a similar methodology to estimate the impulse responses of macroeconomic variables to a government expenditure shock and compare those to impulse responses implied by a standard neoclassic model. The results indicate that hours worked rise, investment shortly increases, while real wages and consumption decrease. Thus they conclude that the standard neoclassic model can account reasonably well for the effects of fiscal policy shocks. A similar conclusion is drawn by Edelberg et al. (1999), who modify a neoclassic growth model distinguishing two types of capital, nonresidential and residential capital.

A structural VAR approach is chosen by Blanchard and Perotti (2002) to identify a government expenditure shock. They require fiscal policy variables not to respond immediately to other innovations in the economy, i.e. they employ the recursiveness assumption. Their findings corroborate the results of Ramey and Shapiro (1998) concerning output and hours worked, but are contradictious with respect to consumption and real wages. Mountford and Uhlig (2005) also use a structural VAR, but do not consider any timing restriction. Instead they employ sign restrictions to restrict the responses of fiscal variables, while the responses of other macroeconomic variables are left open. Besides the different methodology, they additionally allow for a pre-announcement of fiscal policy shocks. Indeed, as it is widely acknowledged and mentioned, most fiscal policy shocks are pre-announced. Their findings, however, confirm the results of Blanchard and Perotti (2002) except for consumption, which only shows a weak positive response.

The debate about the empirical evidence was reopened by Ramey (2008)[2]. Her paper takes up two issues. First, she stresses the importance of the composition of government expenditures. The dataset used by Blanchard and Perotti (2002) includes government consumption as well as government investment expenditures. An increase in the latter can be productive and potentially complement private consumption and investment and therefore lead to a positive response of those variables. For these reasons Ramey advocates to use defense spending as a proxy for government expenditures in the VAR. Second, it states that the findings of the studies differ due to pre-announcement effects, implying that Blanchard and Perotti (2002) employ a faulty timing to identify the fiscal policy shock. In her paper, a neoclassic DSGE model including a pre-announced government expenditure shock is set up and used to sim-

[2]The first version dates back to 2006.

ulate artificial data. It is then demonstrated that, if the pre-announcement of the shock is taken into account, a negative response of consumption is estimated. If not, consumption appears to react positively, a clearly misleading result.

In his summary and discussion of the recent literature, Perotti (2007) acknowledges the concerns with respect to the structural VAR methodology. As a possibility to overcome its weaknesses he suggests to employ annual data and to distinguish between shocks to defense spending and to civilian government spending. However, using annual data, the recursiveness assumption that the fiscal sector does not react contemporarily on the state of the economy might not hold anymore. But, as Perotti mentions, the narrative approach has considerable weaknesses on its own: First, it suffers from a small sample size, second, it is not entirely clear whether the whole change in government expenditures is announced at once or whether it was a combination of small changes, i.e. whether there were numerous revisions of the military budget, occurring one after the other, causing private consumption to respond multiple times.

In Ravn et al. (2007) the authors dismiss Ramey's critique towards the usage of structural VAR models. They point out that shocks are by assumption orthogonal to the information set and consequently identify a structural VAR as in Blanchard and Perotti (2002). However, two papers by Leeper and coauthors, which are concerned with the mapping of estimated reduced form shocks of government expenditures into structural innovations, put this notion into question. In Chung and Leeper (2007) the authors discuss the importance of the intertemporal government budget constraint for a structural VAR analysis. In order to estimate reduced form shocks that can be mapped into structural innovations government debt and private investment should be included into a VAR. Leeper, Walker, and Yang (2008) address the issue of identifying pre-announced tax shocks. They show that due to a difference in the information set of the agents in the economy and the information set of the econometrician aiming at estimating the effects of pre-announced tax shocks, the estimated impulse response functions are biased.

In chapter 3 I estimate a structural VAR. I therefore do not encounter the problems of the narrative approach. I resolve the problem of faulty timing assumptions by not employing a recursive identification scheme, but by taking the pre-announced nature of the shock explicitly into account and restrict the signs of key variables like investment while leaving open the response of the variables of interest. The restrictions are derived from a DSGE model exhibiting forward looking agents. The criticism of Leeper et al. (2008) concerning the estimation of structural VAR is taken into account by imposing the restrictions directly on the impulse response functions of the VAR.

This formulation of a prior distribution on the impulse response functions of the VAR, i.e. requiring them to be in line with the impulse response functions of the DSGE model with forward looking agents, aims at closing the difference in the information sets of the econometrician and agents of the economy.

1.2.5 Robust monetary policy

This section reviews the literature on how policy should be conducted if the policy maker is uncertain about the true economic environment, i.e. when the true model is not known to him.

In his seminal article Brainard (1967) investigated how monetary policy should be conducted if the policy authority is uncertain about the parameters of the model. He finds that in this case, optimal policy should react more cautious. Mccallum (1988) studied the performance of policy rules across different estimated models. He simulated data from each model with the historic estimated policy rule and alternative policy rules under consideration. As a measure how well the policy rule performs, he compared paths for the nominal gdp and inflation and judged them according to their smoothness. He also finds that the policy maker should take model uncertainty into account and determines characteristics of a policy rule performing well across different models.

A first rigorous treatment of optimal policy robust towards model uncertainty was laid out by Hansen and Sargent (2001a), Hansen and Sargent (2001b) and Hansen and Sargent (2003). [3] In their work a benchmark model is formulated and model uncertainty is modeled by an additional error term. This formulation of misspecification results in a set of models, more specifically a set of perturbations of the benchmark model. The set of model perturbations is bounded by assuming that no perturbation can deviate further from the benchmark model than a maximal statistical distance measure. As a robust policy, they define a policy rule which minimizes a loss criterion in the worst case realization of the shock process (minimax-approach).

Subsequently, two issues were discussed and dealt with in the literature. The first issue was the formulation of model misspecification. This does not allow for the assessment of structural model uncertainty, i.e. a discrete set of models considered by the policy maker. Also, deviation from the benchmark, though statistically small, might not be plausible economically. The second issue was the definition of a robust policy rule. Instead of focusing on one realization only, i.e. the worst case scenario, the

[3] In those papers the authors extended an approach already described in Hansen, Sargent, and Tallarini (1999).

whole distribution of outcomes is considered. This approach has been chosen already by Brainard and became more apparent in the literature after Sims (2001) pointed to potential pitfalls of the strategy suggested by Hansen and Sargent.

The results of the studies below often differ depending on the relevant source of structural uncertainty[4] and whether the minimax or the Bayesian approach is employed. It can be generally said that the optimal policy derived from a minimax approach tends to be more aggressive and that parameter uncertainty tends to be less relevant than uncertainty about the (non-nested) structure of the economy.

J. Tetlow and von zur Muehlen (2001) compare a robust policy design with structured and unstructured uncertainty. They find that optimal policy rules concerned with structural model uncertainty are less aggressive than policy rules under unstructured uncertainty and that they are a good approximation of an estimated policy rule. J. Tetlow and von zur Muehlen (2001) also interpret structured uncertainty as the policy maker being uncertain about structural parameters of the model. Soderstrom (2002) follows their interpretation but employs a Bayesian approach.

Levin and Williams (2003) analyze structural uncertainty as uncertainty about a given, discrete set of models using Bayesian and minimax approach. They provide evidence that a policy rule which is robust in the neighborhood of one reference model performs poorly once the models differ more substantially. In Levin, Wieland, and Williams (2003), the authors choose a Bayesian approach to derive a robust policy rule across a set of five discrete models to investigate whether policy rules should be based on forecasts rather than on outcomes. While in those papers the probability of each model is chosen freely, Brock, Durlauf, and West (2005) consider a smaller set of competing theories, but determine the model probabilities by estimating the posterior probability of each model.

Onatski and Stock (2002) compare the performance of robust policy rules derived via the minimax or Bayesian approach considering different structure types of uncertainty, e.g. time-invariant perturbations and time-varying perturbations. Onatski and Williams (2003) investigate optimal policy under parameter and model specification uncertainty (lag length or error term properties) jointly and separately for a backward looking model.

All those studies consider backward looking models only. The succeeding research

[4]In this section I concentrate on the discussion of the sources of uncertainty examined in chapter 4: uncertainty about the parameters of the model and the specification of the model. One strand of the literature stresses another source of uncertainty: data uncertainty. Among those, the most influential studies are Rudebusch (2002), Coenen, Levin, and Wieland (2005), Svensson and Woodford (2004) and Orphanides and Williams (2002).

built on these methodologies but aimed at incorporating forward looking models for a mainly two reasons. First to be more in comply with the Lucas' critique and second in models with a specified utility function and rational expectations, the criterion function employed can be derived from the utility function of the household.

Levin, Onatski, Williams, and Williams (2005) examine a medium scale DSGE model as a benchmark model and additional frictions and features as perturbations. This allows them to derive a micro founded loss criterion, i.e. the households' unconditional expected utility is used as a welfare measure instead of an ad hoc loss function. Employing a Bayesian approach they also consider parameter and model specification uncertainty (the added frictions and features), but do not compute a policy rule which is robust to the source of model specification uncertainty.

Küster and Wieland (2005) combine the minimax and the Bayesian approach and derive a policy rule which is robust to model specification uncertainty, but do not consider a micro-founded loss function. Giannoni (2007) uses a New Keynesian model and the minimax approach to assess parameter uncertainty. He shows that the optimal rule is likely to be more aggressive under parameter uncertainty. Coenen (2007) builds on the work of Levin et al. (2003) and analyzes optimal policy if there are two models under consideration implying a different degree of inflation persistence.

Levine, McAdam, Pearlman, and Pierse (2008) also use a medium-scale New Keynesian model to assess the importance of uncertainty about the degree of indexation in wages and prices on the optimal conduct of policy for the Euro area. They compute optimal simple rules that are robust to this source of specification uncertainty and find similar to Levin et al. (2005) that monetary policy should respond to wage inflation.

Another branch of the literature inspects the optimal policy under model uncertainty problem from a different angle. Cogley, Colacito, and Sargent (2007) and Cogley, Colacito, Hansen, and Sargent (2008) evaluate a setup where the central bank faces uncertainty about two competing aggregate macro models of which one is assumed to be the true data generating process. The central bank seeks to maximize a quadratic loss function, which is weighted with the two model probabilities. To serve this final goal, the policy maker may employ its policy instrument to experiment, to learn and therefore to update its belief about the true model over time. By experimenting systematically the central banker learns faster about the true model and reduces losses due to model uncertainty - even if this leads to transitory suboptimal policies. The authors find it is optimal for the policy maker to pursue this avenue.

Wieland (2000) discuss how the central bank should learn optimally if is is confronted with parameter uncertainty. Orphanides and Williams (2007) analyze optimal policy if the agents posses imperfect knowledge about the economy and learn themselves.

In chapter 4 we employ a Bayesian approach to determine optimal monetary policy under model uncertainty. We choose to model structured model uncertainty by examining a discrete set of models which consists of a benchmark New Keynesian model and reasonable perturbations. Our loss function is micro-founded, i.e. it maximizes households' expected utility. We consider parameter uncertainty and model specification separately and jointly. In contrast to the literature on learning, we do not take a stand on the true data generating process. Since uncertainty about the true model cannot be completely resolved by learning, we focus on how to conduct optimal policy when uncertainty about the true model is persistent.

Chapter 4 is related to Levin et al. (2005) and to Levine et al. (2008). While both work with a medium scale model, our benchmark is a stripped to bare bones New Keynesian model. This modeling and estimation strategy allows us to quantify the importance of each model component for explaining the data and for the optimal conduct of monetary policy separately.

1.3 Outline of the thesis

Chapter 2 addresses the issue of how to identify the structural shocks of a Vector Autoregression (VAR) and how to estimate a dynamic stochastic general equilibrium (DSGE) model when it is not assumed to replicate the data generating process jointly. It proposes a framework to estimate the parameters of the VAR model and the DSGE model: the VAR model is identified by sign restrictions derived from the DSGE model; the DSGE model is estimated by matching the corresponding impulse response functions.

Chapter 3 investigates the effect of a government expenditure shock on private consumption and real wages. A Vector Autoregression is identified by sign restrictions which are in turn derived from a dynamic stochastic general equilibrium (DSGE) model. This allows explicitly to model pre-announcement of a government expenditure shock and its consequences on other economic variables. Since the sign restrictions are not unique across the parameter space of the DSGE model, the DSGE model is estimated by matching the corresponding impulse response functions of the VAR model. Thus the VAR model and the DSGE model are estimated jointly. The results show a significant positive response of private consumption and positive, though not significant, response of real wages.

In chapter 4 we assess the relevant sources of uncertainty for the optimal conduct of monetary policy within (parameter uncertainty) and across a set of nested models (specification uncertainty) using EU 13 data. As our main result, we find that running optimal policy according to the model including all features and frictions does not guard against model uncertainty. Parameter uncertainty matters only if the zero bound on interest rates is explicitly taken into account. In any case, optimal monetary policy is highly sensitive with respect to specification uncertainty implying substantial welfare gains of a optimal rule that incorporates this risk.

Chapter 2

Matching Theory and Data: Bayesian Vector Autoregression and Dynamic Stochastic General Equilibrium Models

This chapter shows how to identify the structural shocks of a Vector Autoregression (VAR) model while simultaneously estimating a dynamic stochastic general equilibrium (DSGE) model that is not assumed to replicate the data-generating process. It proposes a framework for estimating the parameters of the VAR model and the DSGE model jointly: the VAR model is identified by sign restrictions derived from the DSGE model; the DSGE model is estimated by matching the corresponding impulse response functions.

2.1 Introduction

How can we estimate the effects of an exogenous disturbance on the economy? In recent years, two methodologies have become popular to answer this question: the Vector Autoregression (VAR) model and the Dynamic Stochastic General Equilibrium (DSGE) model approach. Both have considerable advantages but also substantial shortcomings. While on the one hand the VAR model is straightforward to estimate, structural shocks cannot be recovered without additional assumptions. The DSGE model, on the other hand, is of a structural form, i.e. it exhibits structural shocks, but it is difficult to determine its parameter values. In this chapter I present a methodology for estimating the effects of exogenous disturbances that combines the advantages of both approaches while overcoming their respective limitations.

I suggest identifying the VAR model with the help of the structural impulse response functions of the DSGE model. Deriving the identifying restrictions from the DSGE model ensures consistency of the identification of the VAR model with the implied structural form of the DSGE model. Moreover, this approach allows the researcher to lay out the assumptions underlying the identification of the VAR model explicitly in the DSGE model and it enables her to include the different assumptions she wants to discriminate between in the DSGE model. In this case a larger class of identifying assumptions can be considered a priori and evaluated afterwards. At the same time, the parameters of the DSGE model are estimated using information from the VAR model. This has the advantage that the DSGE model does not have to be assumed to represent the data-generating process nor to be fully stochastically specified. Therefore, it need not exhibit as many structural shocks as there are observable variables to be explained. Moreover, features and frictions which are not pertinent to the question being examined can be ignored.

More precisely, the VAR model is identified using sign restrictions derived from the structural impulse response functions of the DSGE model, while the DSGE model is estimated by matching the corresponding impulse response functions. Transferring the restrictions via sign restrictions is straightforward and easy to handle: for a given parametrization of the DSGE model the signs of the impulse response functions of the DSGE model define the restrictions for identifying the VAR model. Furthermore, when using sign restriction it is not necessary for the complete number of structural shocks of the VAR model to be identified, nor need the number of structural shocks of the DSGE model correspond to the number of observable variables (variables in the VAR model). The parameter vector of the DSGE model is in turn estimated by matching the corresponding impulse response functions of the VAR model. Thus, it only needs to represent the dynamics of the economy, not the complete data-generating process. Consequently, features and lags which would otherwise have been included to match outliers in the data, but which are not essential to the study, can be dropped.

In order to carry out this estimation procedure, it is necessary to describe the joint distribution of the VAR model and the DSGE model. This chapter presents a methodology for doing so. The methodology is first illustrated by means of a Monte Carlo experiment and then applied to the data. I employ two different DSGE models in each exercise. This is motivated by the fact that the simple DSGE model used in the Monte Carlo Experiment exhibits different signs in the response of each variable depending on the parametrization, i.e. it is a perfect example, but is too stylized to be estimated. The DSGE model used in the estimation exercise does not exhibit this characteristic. Only the response of one variable, the one under investigation, switches signs across the parameter space. However, it is straightforward to be taken to the data.

More precisely, I simulate data from a fiscal theory of the price level (FTPL) model and re-estimate the parameters of the FTPL model and the impulse response functions of the VAR model. The experiment shows that the true impulse response function is indeed found. The FTPL model serves well for illustrating purposes since it can be reduced to two equations in two variables and two shocks. The signs of both variables vary depending on two parameters only. It is less well suited to bringing it to the data. I therefore estimate a DSGE model recently laid out by Ravn, Schmitt-Grohe, Uribe, and Uuskula (2008) to investigate the response of inflation to a monetary policy shock. This DSGE model suits well, since the response of inflation is either positive or negative depending on its parametrization.

The chapter is organized as follows: The next section briefly reviews the relevant literature. The third section outlines the general framework. The fourth section describes the probability distributions and the algorithm suggested to approximate them in detail. The Monte Carlo Experiment is conducted in section 2.5. Section 2.6 applies the methodology to the data and estimates the deep habit model. The last section concludes.

2.2 Related Literature

After Sims's seminal article (Sims, 1980) VAR models became one of the workhorses in macroeconomics despite the problem of identifying structural shocks. Suggestions for resolving the identification problem in a VAR model are manifold. Excellent surveys have been written by Christiano, Eichenbaum, and Evans (1999) and Rubio-Ramirez, Waggoner, and Zha (2005). The approaches most closely related to the methodology presented here are to identify the VAR model by sign restrictions (Uhlig, 2005a; Faust, 1998) or by probabilistic restrictions (Kociecki, 2005). Identification employing sign restrictions attempts to restrict the signs of the impulse response functions of some variables, while the variable of interest is unrestricted. In Kociecki (2005), a prior distribution for the impulse response functions is formulated and transformed into a prior distribution for the coefficients of the structural VAR model. Both approaches depend on the availability of a priori knowledge on the behavior of some impulse response functions.

With regard to explicitly basing the identifying assumptions on DSGE models, two strands of literature have emerged recently. One derives the identifying assumptions from a DSGE model (Altig et al. (2002), DelNegro and Schorfheide (2004) and Sims (2006b)); the other suggests, once the DSGE model is large enough, estimating the DSGE model and thereby directly infering on the impulse responses (as in Smets and Wouters (2003) and Smets and Wouters (2007)).

Due to advances in computational power, the estimation of DSGE models has lately become very popular. The procedures differ depending on the econometric interpretation of the DSGE model. Geweke (1999a) distinguishes between a strong and weak interpretation. The former requires the DSGE model to provide a full description of the data-generating process. It is the more common one nowadays despite its shortcomings: first, the DSGE model already puts a lot of structure on the impulse responses a priori, i.e. it often does not allow an investigation of the sign of a response and might therefore not be appropriate as a research tool. Second, not all parameters of the DSGE model can be identified (see Canova and Sala (2006) and Beyer and Farmer (2006)). Finally, not all economists might feel comfortable with the assumption that the DSGE model is a proper representation of the data-generating process. Instead, as mentioned in Christiano et al. (2005), the DSGE model is best suited to replicate the implied dynamics in the data, i.e. the impulse response functions. This is the weak econometric interpretation. Following this road Ravn et al. (2007), Mertens and Ravn (2008) as well as Ravn et al. (2008) estimate a DSGE model given the impulse response function of the VAR model by minimizing the distance between the corresponding impulse response functions. In contrast to them I do not consider the impulse response functions of the VAR model as given, i.e. as identified a priori. In the case of timing or long-run restriction the VAR model is identified and considering the impulse response functions as given is justified. This chapter addresses the cases when the identifying restrictions are not a priori clear or when the researcher chooses to use sign restrictions. Sign restrictions derived from a DSGE model will only in very rare cases be unique across the parameter space of the DSGE model. In those cases the impulse response functions are not identified and one cannot proceed as for instance in Ravn et al. (2007), Mertens and Ravn (2008) or Ravn et al. (2008).

The methodology presented in this chapter is in the spirit of the former strand of the literature, i.e. it bases the identification of the VAR model on restrictions derived from the DSGE model. It differs from the existing literature in the following aspects.

Altig et al. (2002) and DelNegro and Schorfheide (2004) employ the rotation matrix of the DSGE model to identify the VAR model. To do this, the DSGE model has to be fully stochastically specified. In the case of DelNegro and Schorfheide (2004), additional dummy observations derived from the model are used to augment the VAR model as suggested originally by Ingram and Whiteman (1994). While one can control for the prior weight of the dummy observations, one cannot control for the prior weight of the implied dynamics of the DSGE. The methodology proposed here differs from this by not employing the implied rotation matrix of the DSGE model to identify the VAR model, and therefore not requiring the DSGE model to be fully stochastically specified.

Sims (2006b) extends the idea to augment the VAR model with dummy observations in a more general framework. In his framework, the tightness of the prior can be varied across frequencies and the number of structural shocks does not need to equal the number of observations. The main difference to Sims (2006b) is that I suggest employing the implied sign and shape restrictions (as described in Uhlig (2005a)) to identify the VAR model as it is more simple and straightforward to use.

In recent studies, Lanne and Lütkepohl (2005), Lanne and Lütkepohl (2008), and Lanne, Lütkepohl, and Maciejowska (2009) employ additional statistical properties of the error terms to identify the VAR model. Lanne and Lütkepohl (2005) make use of possible non-normal distributions of the error terms and extract additional identifying information from this. Lanne and Lütkepohl (2008) use the insight of Rigobon (2003) that a VAR model can be identified exploiting changes in volatility. Given any exact identifying scheme this characteristics delivers over-identifying restrictions which can be used to test different identification schemes. In Lanne et al. (2009) the authors combine the properties of mixed normal distributions and regime changes in the volatility of the error term and show that the VAR model is just identified, given that the shocks are orthogonal across regimes and only the variances of the shocks change across regimes. The methodology presented in this chapter does not hinge on special properties of the error terms. It applies also in cases where the residuals are normally distributed.

2.3 Framework

In this section I set up the VAR model and its corresponding Vector Moving Average (VMA) representation. The issue of how the structural impulse response can be identified is equivalent for both notations. For every period, the impulse response functions of a VAR model can be expressed solely in terms of the coefficients of the VMA model of that period. Setting up the framework in terms of the VMA representation makes the subsequent analytical calculations less demanding. Since the VAR model is connected with the DSGE model via their implied dynamics, the notation necessary for the DSGE model is introduced before the central idea of how to derive the joint posterior distribution for the VAR model and the DSGE model is presented. Afterwards, the framework is related to existing and nested approaches.

2.3.1 The VAR model and its corresponding VMA model

The structural VAR model containing m variables is defined as:

$$A^{-1}Y_t = A_1Y_{t-1} + A_2Y_{t-2} + \dots A_lY_{t-l} + \epsilon_t, t = 1, \dots, T \tag{2.1}$$

Y_t is a $m \times 1$ vector at date $t = 1 + l, \dots, T$, A and A_i are coefficient matrices of size $m \times m$ and ϵ an *i.i.d.* one-step-ahead forecast error, distributed: $\epsilon \sim \mathcal{N}(0, I_{m\times m})$.

The impulse response function φ^V of the VAR model is defined as the response of Y to an innovation in ϵ. Denote the VMA representation as:

$$Y_t = \sum_{i=0}^{\infty} \Theta_i \epsilon_{t-i}, \tag{2.2}$$

where Θ_i denotes a moving average coefficient matrix. The impulse response function of a VAR model to an innovation in variable i at horizon k φ^V_{jk} can be computed directly as:

$$\varphi^V_{jk} = \Theta_{jk}, \tag{2.3}$$

where i depicts the i-th column. Due to the assumption that $\Sigma_\epsilon = I_{m\times m}$, this structural moving average representation cannot be estimated directly. Instead the reduced form moving average representation with error term $u_t = A\epsilon_t$, where $u \sim \mathcal{N}(0, \Sigma)$, is estimated. The reduced form moving average coefficients are defined as $\Phi_i = \Theta_i A^{-1}$:

$$Y_t = \sum_{i=0}^{\infty} \Phi_i u_{t-i} \tag{2.4}$$

The factorization $\Sigma = A'A$ does not have a unique solution, which leads to an identification problem of A.

It is important to note that any stationary moving average representation can be approximated by a reduced form VAR model, which takes the form:

$$Y_t = B_1Y_{t-1} + B_2Y_{t-2} + \dots B_lY_{t-l} + u_t, t + 1, \dots, T \tag{2.5}$$

with $B_i = AA_i$, $u_t = A\epsilon_t$ and $u \sim \mathcal{N}(0, \Sigma)$. While the framework is set up in terms of VMA representation, it can be easily estimated as a VAR model.

2.3.2 The DSGE model

The fundamental solution of the DSGE model is given by[1]:

$$\hat{x}_t = T(\tilde{\theta})\hat{x}_{t-1} + R(\tilde{\theta})z_t, \tag{2.6}$$

where z is a vector collecting the structural shocks of the DSGE model, while $T(\tilde{\theta})$ and $R(\tilde{\theta})$ are matrices one obtains after solving a DSGE model with standard solution techniques.

The impulse response functions of the variables in x to a structural shock i at horizon k φ^D_{ik} are given by:

$$\varphi^D_{i,0} = R(\tilde{\theta})z_i, k = 0 \tag{2.7}$$

$$\varphi^D_{i,k} = T(\tilde{\theta})\varphi^D_{k-1,i}, k = 1, 2, ...K. \tag{2.8}$$

The vector of structural parameters of the DSGE model defined as in (2.6) does not contain any variances or covariances of a measurement error or any error term emerging from confronting the DSGE model with the data, but only the variances of the structural shocks. When the DSGE model is estimated by matching the corresponding impulse response functions, an additional error term occurs. Its variance covariance matrix is denoted by Ω and is also estimated. The vector comprising the vector of deep parameters $\tilde{\theta}$ and the vectorized Ω is defined as $\theta = [\begin{array}{cc} \tilde{\theta} & vec(\Omega) \end{array}]'$.

2.3.3 The idea in a nutshell

On the one hand, the distribution of the parameters of the DSGE model is estimated by matching the corresponding impulse response function of the VMA model. On the other hand, the distribution of structural impulse response functions of the VMA model are identified by applying sign restrictions which are derived from the DSGE model. Both distributions are therefore conditional distributions: they depend on a realization of the impulse response function of the VMA model and on restrictions from the DSGE model, i.e. a realization of a vector of structural parameters of the

[1] $\widehat{x}_t$ denotes the percentage deviation of the generic variable x_t from a deterministic steady state x chosen as approximation point.

DSGE model, respectively. This section sets out how the conditional distributions can be combined to derive the joint distribution.

The joint posterior distribution of θ and φ, given a matrix with time series observations Y, $p(\theta, \varphi^V|Y)$, can be decomposed in different ways, depending on whether the DSGE model is employed to identify the VMA model or not. In the latter case the joint posterior is given by:

$$p(\varphi^V, \theta|Y) = p(\varphi^V|Y)p(\theta|\varphi^V). \tag{2.9}$$

This equation can be justified twofold: In the case that the DSGE model is estimated by matching the corresponding impulse response functions and not time series observations, the distribution of θ conditional on φ^V and Y is equal to the distribution of θ conditional on φ^V only[2]. The second justification is shown by Smith (1993) and DelNegro and Schorfheide (2004) and discussed in section 2.3.4, when setting the framework in a broader context.

In the case that the likelihood of the VMA model impulse response functions depends on restrictions from the DSGE model, $p(\theta, \varphi^V|Y)$ is given as:

$$p(\varphi^V, \theta|Y) = p(\varphi^V|\theta, Y)p(\theta|Y). \tag{2.10}$$

The framework presented in this chapter is based on the argument that both distributions are at least proportionally equal:

$$p(\varphi^V|Y)p(\theta|\varphi^V) \propto p(\varphi^V|\theta, Y)p(\theta|Y), \tag{2.11}$$

and can be approximated sufficiently well by Monte Carlo Markov Chain Methods.

Denote the Jacobian matrix collecting the derivatives of φ^V with respect to Φ by $J(\varphi^V \rightarrow A, \Phi)$. Considering the relationship between the coefficients of the VMA model and the impulse response function of the VMA model ($\Phi_i A = \varphi_i^V$):

$$p(\varphi^V|\theta, Y) = p(A, \Phi|\theta, Y)J(\varphi^V \rightarrow A, \Phi), \tag{2.12}$$

[2]It then holds:

$$p(\theta|\varphi^V, Y)p(\varphi^V|Y) = p(\varphi^V|Y)p(\theta|\varphi^V)$$

equation (2.11) is given by:

$$p(\varphi^V|Y)p(\theta|\varphi^V) \propto p(A,\Phi|\theta,Y)J(\varphi^V \to A,\Phi)p(\theta|Y). \tag{2.13}$$

Note that the conditional distributions of interest ($p(\theta|\varphi^V)$ and $p(A,\Phi|\theta,Y)J(\varphi^V \to A,\Phi)$ are on different sides of the proportionally sign in (2.13). It is therefore possible to employ a Gibbs sampling algorithm, i.e. to draw from two conditional distributions in order to evaluate the joint distribution. In the following section I will first relate the approach to existing methodologies before I describe both conditional distributions in detail.

2.3.4 Nested approaches

Taking a broader perspective, several closely-related methodologies evolve as special cases of this approach: the pure sign restriction approach of Uhlig (2005a), the DSGE-VAR methodology of DelNegro and Schorfheide (2004) and the case of probabilistic restrictions of Kociecki (2005).

The latter arises if the restrictions derived from the DSGE model are constant across the parameter space. Then it is possible to generate a prior distribution for the impulse response functions of the VAR model from the DSGE model and use it as a prior for the parameters of the VAR model. Since, as pointed out by Kociecki (2005), the sign restriction approach is a special case of the probabilistic approach, this methodology is also nested. The sign restriction approach arises if the prior distribution for some impulse response function exhibits a very small variance, i.e. determines the sign of this impulse. It is equivalent to using an indicator function placing zero probability weight on VAR model parameter regions whenever the a priori sign restrictions are not satisfied. Therefore, in the case that the DSGE model determines constant sign restrictions across the parameter space it is not necessary to draw from the conditional distribution of θ. One only needs to draw from $p(A,B|\theta,Y)$.

The DSGE-VAR methodology arises once the framework is rewritten in terms of the parameters instead of the impulse response functions of the VAR model, and in the case that the DSGE model is fully stochastically specified.

$$p(A,B|Y)p(\theta|A,B) \propto p(A,B|\theta,Y)p(\theta|Y) \tag{2.14}$$

The right-hand side is the expression used to evaluate the joint posterior distribution of $p(A,B,\theta|Y)$: since the DSGE model is fully stochastically specified it is possible to derive an analytical solution for the marginal posterior of θ. The decomposition on the left-hand side again legitimates the decomposition used in (2.9): the posterior distribution of the parameters of the VAR does not depend on the vector of structural parameters of the DSGE model. As argued in DelNegro and Schorfheide (2004) and Smith (1993), A and B can then be used to learn about the parameter vector θ.

2.4 Evaluating the joint distribution

In this section the conditional distributions employed in the estimation process are described in detail. I start by describing the distribution of the VMA model conditional on the parameter vector of the DSGE model. Then the distributions of the parameters of the DSGE conditional on the impulse response functions of the VMA model are set out.

2.4.1 Conditional distribution of the VMA model parameters

The conditional distribution described in this section is $p(\varphi^V|\theta, Y)$ from the right-hand side of (2.11). It is conditional since the prior distribution for the impulse response functions $p(\varphi^V) = p(\varphi^V|\theta)$ is derived from the DSGE model[3] The posterior distribution of the structural impulse responses φ^V is obtained by combining the coefficient estimates of the reduced-form VMA model Φ with an impulse matrix A. In order to write this distribution in terms of the reduced-form VMA model coefficients and the impulse matrix, it has to be scaled by the Jacobian $J(\varphi^V \to A, \Phi)$. The prior distribution for the structural impulse response function is set out and the Jacobian is derived in the first subsection.

Afterwards, the distribution $p(A, \Phi|\theta)$ is decomposed into a conditional distribution of the VMA model coefficients and a marginal distribution of the impulse matrix A:

$$p(\Phi, A|\theta) = p(\Phi|A, \theta)p(A|\theta). \tag{2.15}$$

Combining this prior distribution with the likelihood yields the posterior distribution $p(A, \Phi|\theta, Y)J(\varphi^V \to A, \Phi)$. The likelihood is described in the third part of this section. The resulting posterior distribution is difficult to evaluate for various reasons: given the restrictions, it is, to the best of my knowledge, not possible to draw the impulse matrix A of the VMA model for a reasonably large set of variables. It is not possible

[3] The impulse response functions of the DSGE model define a probability distribution of impulse response functions dependent on θ.

if only submatrices, i.e. fewer shocks than variables, are considered. This also implies that the distributions conditional on A are not defined, causing problems in the case that the restrictions are formulated for more than one period.

I therefore suggest in the fourth section deriving the restrictions from the DSGE model as sign restrictions. For each realization of the impulse response function of the DSGE model the corresponding sign restrictions are put on the VMA model. The coefficients of the VMA model are conditional on the impulse response functions of the DSGE model, similar to Uhlig (2005a), where the posterior distribution of the VAR parameters is multiplied by an indicator function that puts zero probability in parameter regions whenever the restrictions derived from the DSGE model are not satisfied. The distribution of parameters of the DSGE model θ defines a set of restrictions put on the parameters of the VMA model. This conditional prior distribution combined with the likelihood then yields the posterior distribution. A further simplification is considered in the concluding part of this section: the approximation of the VMA model by a VAR model.

2.4.1.1 The Jacobian $J(\varphi^V \to A, \Phi)$

Denote the impulse response functions in period k as φ_k^V. If all shocks are included, the matrix is of size $m \times m$, where the entry i, j corresponds to the response of variable i to an innovation in variable j. The prior for the impulse responses has to be specified for as many periods as there are impulse response functions to be estimated. The vectorized impulse responses are assumed to be normally distributed:

$$\begin{bmatrix} vec(\varphi_0) \\ vec(\varphi_1) \\ vec(\varphi_2) \\ \vdots \\ vec(\varphi_l) \end{bmatrix} \sim \mathcal{N}\left(\begin{bmatrix} vec(\bar{\varphi}_0) \\ vec(\bar{\varphi}_1) \\ vec(\bar{\varphi}_2) \\ \vdots \\ vec(\bar{\varphi}_k) \end{bmatrix}, \begin{bmatrix} \bar{V}_{00} & \bar{V}_{01} & \bar{V}_{02} & \cdots & \bar{V}_{0l} \\ \bar{V}_{10} & \bar{V}_{11} & \bar{V}_{12} & \cdots & \bar{V}_{1k} \\ \bar{V}_{20} & \bar{V}_{21} & \bar{V}_{22} & \cdots & \bar{V}_{2k} \\ \vdots & & & & \\ \bar{V}_{k0} & \bar{V}_{k1} & \bar{V}_{k2} & \cdots & \bar{V}_{kk} \end{bmatrix} \right). \tag{2.16}$$

The probability distribution $p(\varphi_0, \varphi_1, ..., \varphi_k)$ can be decomposed into a marginal distribution of $p(\varphi_0)$ and succeeding conditional distributions:

$$p(\varphi_0, \varphi_1, ..., \varphi_k) = p(\varphi_k|\varphi_{k-1} \cdots \varphi_0)p(\varphi_{k-1}|\varphi_{k-2} \cdots \varphi_0) \cdots p(\varphi_1|\varphi_0)p(\varphi_0), \tag{2.17}$$

with

$$p(vec(\varphi_0)) = \mathcal{N}(vec(\bar{\varphi}_0), \bar{\bar{V}}_{00}) \tag{2.18}$$

$$p(vec(\varphi_i|vec(\varphi_{i-1})\cdots vec(\varphi_0)) = \mathcal{N}(\theta_i, \Delta_{ii}), i = 1\cdots k, \tag{2.19}$$

and θ_i and Δ_{ii} abbreviate the usual definitions for conditional distributions:

$$\theta_i = vec(\bar{\varphi}) + \left[\begin{array}{c} \bar{V}_{i0}\cdots\bar{V}_{ii-1} \end{array}\right] \left[\begin{array}{ccc} \bar{V}_{00} & \cdots & \bar{V}_{0i-1} \\ \vdots & \ddots & \vdots \\ \bar{V}_{i-1,0} & \cdots & \bar{V}_{i-1,i-1} \end{array}\right]^{-1} \left[\begin{array}{c} vec(\varphi_0 - \bar{\varphi}_0) \\ \vdots \\ vec(\varphi_{i-1} - \bar{\varphi}_{i-1}) \end{array}\right]$$

$$\Delta_{ii} = \bar{V}_{ii} - \left[\begin{array}{ccc} \bar{V}_{i0} & \cdots & \bar{V}_{ii-1} \end{array}\right] \left[\begin{array}{ccc} \bar{V}_{00} & \cdots & \bar{V}_{0i-1} \\ \vdots & \ddots & \vdots \\ \bar{V}_{i-1,0} & \cdots & \bar{V}_{i-1,i-1} \end{array}\right]^{-1} \left[\begin{array}{c} \bar{V}_{0i} \\ \vdots \\ \bar{V}_{i-1,i} \end{array}\right]$$

In order to write the prior distribution in terms of the reduced form coefficients it is necessary to scale the probability distribution with the Jacobian:

$$p(\varphi) = p(f(\Phi))J(\varphi \Rightarrow \Phi). \tag{2.20}$$

The relationship between structural and reduced form moving average coefficients is given by:

$$\varphi_0 = A \tag{2.21}$$

$$\varphi_i = \Phi_i A, i = 1\cdots k \tag{2.22}$$

.

Note that I have left out Φ_0 since this matrix is normalized to an identity matrix by assumption. This also indicates that it is not possible to infer on φ_0 from the estimated reduced VMA model.

The Jacobian is calculated in the following way. Applying the vec-operator yields:[4]

$$vec(\varphi_i) = (A' \otimes I_{m\times m})vec(\Phi_i).$$

[4]Note that $vec(AB) = (I \otimes A)vec(B) = (B' \otimes I)vec(A))$

The Jacobian matrix is defined as:

$$J(\varphi \to \Phi) = det \begin{bmatrix} \frac{\partial vec(\varphi_1)}{\partial vec(\Phi_0)} & \frac{\partial vec(\varphi_1)}{\partial vec(\Phi_1)} & \cdots & \frac{\partial vec(\varphi_1)}{\partial vec(\Phi_k)} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial vec(\varphi_k)}{\partial vec(\Phi_0)} & \frac{\partial vec(\varphi_k)}{\partial vec(\Phi_1)} & \cdots & \frac{\partial vec(\varphi_k)}{\partial vec(\Phi_k)} \end{bmatrix}. \tag{2.23}$$

Due to the fact that $\frac{\partial vec(\varphi_i)}{\partial vec(\Phi_j)} = 0$ for $j > i$, the matrix becomes a block triangular matrix and the determinant is given by:

$$J(\varphi \to \Phi) = |\frac{\partial vec(\varphi_0)}{\partial vec(\Phi_0)}| \times |\frac{\partial vec(\varphi_1)}{\partial vec(\Phi_1)}| \cdots |\frac{\partial vec(\varphi_k)}{\partial vec(\Phi_k)}|$$

$$J(\varphi \to \Phi) = |(A' \otimes I_{m\times m})|^k = |A|^{mk} \tag{2.24}$$

$$\tag{2.25}$$

2.4.1.2 Decomposition of the distribution $p(\Phi, A) = p(\Phi|A)p(A)$

A prior distribution for the reduced form coefficients conditional on $\varphi_0 = A$ is formulated as:

$$p(A, \Phi_1, ..., \Phi_k) = p(\Phi_k|\Phi_{k-1} \cdots \Phi_0)p(\Phi_{k-1}|\Phi_{k-2} \cdots A) \cdots p(\varphi_1|A)p(A)J(\varphi \to \Phi), \tag{2.26}$$

where

$$p(vec(A)) = \mathcal{N}(vec(\bar{\varphi}_0), \bar{\bar{V}}_{00}) \tag{2.27}$$

$$p(vec(\Phi_i)|vec(\Phi_{i-1}) \cdots vec(A)) = \mathcal{N}(\bar{\Phi}_i, \bar{\bar{V}}_i i) \tag{2.28}$$

with

$$\bar{\Phi}_i = (A' \otimes I_{m\times m})\theta_i \tag{2.29}$$

$$\bar{\bar{V}}_i i = (A^{-1'} \otimes I_{m\times m})\Delta_{ii}(A^{-1'} \otimes I_{m\times m}). \tag{2.30}$$

2.4.1.3 The likelihood for the reduced-form coefficients

Consider the VMA(k) process:

$$Y_t = u_t + \Phi_1 u_{t-1} + \Phi_2 u_{t-2} + \cdots + \Phi_k u_{t-k}. \tag{2.31}$$

This can be written in state space form:

$$\xi_{t+1} = F\xi_t + U_{t+1} \tag{2.32}$$

$$y_t = H\xi_t, \tag{2.33}$$

where

$$\underset{m*k\times 1}{\xi_t = \begin{bmatrix} u_t & \cdots & u_{t-k} \end{bmatrix}'}$$

$$F = \underset{m*k\times m*k}{\begin{bmatrix} 0 & 0 & \cdots & 0 & 0 \\ I_m & 0 & \cdots & 0 & 0 \\ 0 & I_m & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & I_m & 0 \end{bmatrix}}$$

$$U_{t+1} = \underset{m*k\times 1}{\begin{bmatrix} u_{t+1} & 0 & \cdots & 0 \end{bmatrix}'}$$

$$H = \underset{m*\times m*k}{\begin{bmatrix} I_m & \Phi_1 & \cdots & \Phi_k \end{bmatrix}}.$$

Given an initial condition for y_0 and Σ_0, the likelihood can then be written as:

$$p(y_T, \dots, y_0|\Phi_1, ..., \Phi_k, \Sigma) = p(y_T|y_{T-1} \dots y_0, \Phi_1, ..., \Phi_k, \Sigma) \cdots p(y_0|\Phi_1, ..., \Phi_k, \Sigma), \tag{2.34}$$

where:

$$p(y_t|y_{t-1} \dots y_0, \Phi_1, ..., \Phi_k) = \mathcal{N}(y_{t|t-1}, \Sigma_{t|t-1}). \tag{2.35}$$

$y_{t|t-1}$ and $\Sigma_{t|t-1}$ denote the optimal forecast at time t, which is a function of the coefficient matrices. The impulse matrix A is not part of the likelihood function, instead the variance covariance matrix $\Sigma = A'A$.

2.4.1.4 The posterior distribution

The posterior of the reduced form coefficients is derived by combining (2.34) and (2.26):

$$p(\Phi_1, \cdots \Phi_k, A|\theta, Y) = p(y_T, \dots, y_0|\Phi_1, ..., \Phi_k, \Sigma)p(A, \Phi_1, ..., \Phi_k|\theta). \tag{2.36}$$

To identify the impulse matrix A from the likelihood estimate of the variance covariance matrix I utilize the prior distribution for $\varphi_0 = A$ derived from the DSGE model in the following way: the impulse matrix $\breve{A}$ is defined as a sub matrix of A of size $m \times n$ where

n is the number of structural shocks under consideration, i.e. the structural shock of interest as well as other shocks necessary to distinguish this shock. These shocks have to be included in the DSGE model as well. In order to indicate that the restrictions put on A rely on the model and therefore its parameter vector θ, I write $\breve{A}(\theta)$. Given a number of rowvectors q_j forming an orthonormal matrix Q and the lower Cholesky decomposition of Σ, $\tilde{A}$, $\breve{A}(\theta)$ is defined as: $\breve{A}(\theta) = \tilde{A}Q(\theta)$.

Every realization of the vector of the parameters of the DSGE model θ is associated with an impulse response function of the DSGE model and a realization of $\breve{A}(\theta)$. A sequence of realizations of θ yields a sequence of restrictions and therefore a related prior probability distribution. Given a realization of an impulse response function of the DSGE model φ^D the posterior distribution is evaluated the following way:

1. Derive the sign restrictions from φ^D.
2. Draw a realization of Φ and Σ from the distribution (2.36).
3. Calculate $\tilde{A}$ and draw $Q(\theta)$ from a uniform distribution such that $\breve{A}(\theta) = \tilde{A}Q(\theta)$ fulfils the sign restriction.
4. Given A, compute the structural impulse responses from $\varphi_i = \Phi_i A, i = 1 \cdots k$.

2.4.1.5 The conditional distribution of the VAR model

Estimating a VAR model is less complicated. In practice whenever possible, i.e. if the VMA model is stationary, it is approximated by a VAR model[5]. In this section I therefore briefly lay out the approach for this case.

As shown by Uhlig (1997), the prior distribution for B and Σ can be specified choosing appropriate B_0, N_0, S_0, v_0 as:

$$vec(B)|\Sigma \sim \mathcal{N}(vec(B_0), \Sigma \otimes N_0^{-1}) \tag{2.37}$$
$$\Sigma \sim \mathcal{IW}(v_0 S_0, v_0). \tag{2.38}$$

Denote the maximum likelihood estimates of Σ and B as $\tilde{\Sigma} = \frac{1}{T}(Y - X\hat{B})'(Y - X\hat{B})$

[5]I will employ the expression VAR model in the remaining sections too.

and $\hat{B} = (X'X)^{-1}X'Y$. The posterior is then given as[6]:

$$vec(B)|\Sigma \sim \mathcal{N}(vec(B_T), \Sigma \otimes N_T^{-1}) \tag{2.39}$$
$$\Sigma \sim \mathcal{IW}(v_T S_T, v_T), \tag{2.40}$$

where

$$N_T = N_0 + X'X \tag{2.41}$$
$$B_T = N_T^{-1}(N_0 B_0 + X'X\hat{B}) \tag{2.42}$$
$$S_T = \frac{v_0}{v_T} S_0 + \frac{T}{v_T}\tilde{\Sigma} - \frac{1}{v_T}(B_0 - \hat{B})' N_0 N_T^{-1} X'X(B_0 - \hat{B}) \tag{2.43}$$
$$v_T = v_0 + T. \tag{2.44}$$

Drawing from a joint posterior of B, Σ and $\breve{A}(\theta)$ is conducted in the following steps:

1. The impulse responses of the DSGE determine the restrictions put on $\breve{A}(\theta)$.
2. Draw B and Σ from the posterior (2.39) and (2.40).
3. Calculate $\tilde{A}$ and draw $Q(\theta)$ from a uniform distribution such that $\breve{A}(\theta) = \tilde{A}Q(\theta)$ fulfills the sign restriction.

2.4.2 The conditional distribution of the DSGE model parameters

Since the DSGE model is not assumed to be a proper representation of the data-generating process, the structural parameters are not estimated by matching the data Y. Instead, the DSGE model is assumed to replicate the implied dynamics of the data, i.e. the impulse response functions of the VAR model. This induces matching a given realization of the impulse response function of the VAR model to the i-th shock at horizon k, $\varphi_{i,k}^V$:

$$\varphi_{i,k}^V = \varphi_{i,k}^D(\tilde{\theta}) + \omega_{i,k}. \tag{2.45}$$

Stacking the impulse response functions over $1, .., K$ periods together yields:

$$\varphi_i^V = \varphi_i^D(\tilde{\theta}) + \omega_i \tag{2.46}$$

[6] A formal derivation is given in appendix A.1

with all vectors of dimension $m{*}k{\times}1$. The error term ω_i has the property $E[\omega_i\omega_i'] = \Omega_{\omega_i}$, which is part of the vector θ. Since the structural shocks are assumed to be independent, the probability of $p(\theta|\varphi^V)$ can be written as:

$$p(\theta|\varphi^V) = p(\theta|\varphi_1^V, \varphi_2^V, \cdots \varphi_i^V) = p(\theta|\varphi_1^V)p(\theta|\varphi_2^V)\cdots p(\theta|\varphi_i^V). \tag{2.47}$$

The vector θ is estimated in two steps: First Ω_{ω_i} is estimated, and afterwards the vector of deep parameters $\tilde{\theta}$. The variance covariance is estimated by making use of the relationship:

$$\omega_i = \varphi_i^V - \varphi_i^D(\tilde{\theta}). \tag{2.48}$$

For every realization of φ_i^V a reasonable number of draws from $p(\theta)$ is taken[7], and the corresponding impulse response function $\varphi_i^D(\tilde{\theta})$ and the error terms are computed. $\tilde{\Omega}_{\omega_i}$ is then estimated as the covariance matrix of these error terms. For each shock i the likelihood $l_i\left(\tilde{\theta}|\varphi_i^V, \tilde{\Omega}_{\omega_i}\right)$is given by:

$$l_i\left(\tilde{\theta}|\varphi_i^V, \tilde{\Omega}_{\omega_i}\right) = -\frac{Km}{2}ln(2\pi) - \frac{1}{2}ln(|\tilde{\Omega}_{\omega_i} \otimes I_K|) - \frac{1}{2}(\omega_i)'\left(\tilde{\Omega}_{\omega_i} \otimes I_K\right)^{-1}(\omega_i). \tag{2.49}$$

Combining this likelihood with a prior distribution for θ yields the posterior distribution.

One potential issue arising when matching impulse response functions of a DSGE model and a VAR model was pointed out by McGrattan, Chari, and Kehoe (2005): the implied VAR model representation of the DSGE model might be of infinite order but the empirical VAR model is often of lower order. One solution, suggested by Cogley and Nason (1995), is to simulate artificial time series from the DSGE model, estimate a VAR model from the artificial time series and compare this VAR model to the VAR model estimated from the actual data.

2.4.3 Sampling algorithm for the joint posterior distribution

In order to evaluate the joint posterior distribution of the parameters of the DSGE model and the VAR model I propose a Gibbs sampling algorithm combined with a Metropolis-Hastings step. The Gibbs sampling algorithm allows to draw from the conditional distributions laid out in detail in sections 2.4.1.4 and 2.4.2. The Metropolis-Hastings step is an acceptance/rejection sampling algorithm that determines the prob-

[7] In the simulation and estimation I used 50 draws per realization.

ability space in which the implied impulse response functions of the DSGE model and those of the VAR model coincide. It is carried out I times.

The algorithm can roughly be summarized in the following way. At each iteration $i = 1, \ldots, I$ d-draws are taken from the conditional densities $p(\theta|\varphi^V)^i$ and $p(\varphi^V|\theta, Y)^i$.[8] These draws form candidate distributions $p(\theta|\varphi^V)^{\tilde{i}}$ and $p(\varphi^V|Y, \theta)^{\tilde{i}}$. Via acceptance and rejection, the candidate distributions are compared with $p(\theta|\varphi^V)^i$ and $p(\varphi^V|\theta, Y)^i$. Draws with higher posterior density are kept and form the new densities $p(\theta|\varphi^V)^{i+1}$ and $p(\varphi^V|\theta, Y)^{i+1}$. More precisely, at each iteration $i = 1, \ldots, I$ the algorithm involves the following steps:

1. Draw $j = 1 \ldots d$ times from $p(\theta|\varphi^V)^i$.
2. For every realization θ_j of the vector of deep parameters of the DSGE model derive the corresponding sign restriction.
3. Draw Σ_j from (2.40) and B_j from (2.39). Compute the lower Cholesky decomposition and find an $\breve{A}_j = \tilde{A}_j Q_j$ fulfilling the sign restrictions from $\varphi_j^D(\tilde{\theta}_j)$. Compute φ_j^V, yielding $p(\varphi^V|Y, \theta)^{\tilde{i}}$.
4. For every realization of φ_j^V derived in step 3 find the θ that maximizes (2.49) combined with the prior $p(\theta)$. This yields $p(\theta|\varphi^V)^{\tilde{i}}$.
5. Do acceptance-rejection by comparing $p(\theta|\varphi^V)^{\tilde{i}}$ with $p(\theta|\varphi^V)^{i-1}$. Keep the corresponding vectors from $p(\varphi^V|\theta)^{\tilde{i}}$. This yields $p(\theta|\varphi^V)^{i+1}$ and $p(\varphi^V|\theta)^{i+1}$.
6. Start again at 1.

The chain converges if $p(\theta|\varphi^V)^i$ and $p(\theta|\varphi^V)^{i-1}$ and also $p(\varphi^V|\theta)^i$ and $p(\varphi^V|\theta)^{i-1}$ are similar, i.e. the acceptance rate is low.

In the remaining sections of the chapter I will discuss the properties of the sampling algorithm in more detail using a Monte Carlo experiment, i.e. specify precisely the number of iterations and the convergence of the algorithm. Afterwards I will put the methodology to work and confront it with the data.

2.5 Example 1: A Monte Carlo Experiment

In order to illustrate the methodology suggested above I use a simple fiscal theory of the price level (FTPL) model as described in Leeper (1991) to identify the response of

[8] In the first iteration step $p(\theta|\varphi^V)^1 = p(\theta)$.

inflation to a monetary policy shock, i.e. an unexpected increase in the interest rate. The FTPL model is chosen because it can be reduced to two equations in real debt and inflation. It is the most simple DSGE model exhibiting different signs of the impulse response functions depending on two parameters only. Furthermore, the solution and properties of the FTPL model are well known by economists, which makes the example very transparent.

I simulate data from the FTPL model and using the methodology outlined above show that the 'true' signs of the impulse response functions and the corresponding distribution of the parameters of the FTPL model are found, even if the chain is initialized with a wrong guess.

2.5.1 The FTPL model

The representative household maximizes its utility in consumption [9] c_F and real money balances m_F:

$$U_t = log(c_{F,t}) + log(m_{F,t}) \tag{2.50}$$

subject to the budget constraint:

$$c_{F,t} + m_{F,t} + b_{F,t} + \tau_{F,t} = yF + \frac{1}{\pi_{F,t}} m_{F,t-1} + \frac{R_{F,t-1}}{\pi_{F,t}} b_{F,t}, \tag{2.51}$$

where b_F denotes bond holdings, τ_F lump sum taxes, y_F income, R_F nominal interest rates and π_F inflation. Small letters denote real variables, capital letters nominal variables.

The government has to finance government expenditures g_F by issuing bonds, collecting taxes and seignorage. The budget constraint is therefore given by:

$$b_{F,t} + m_{F,t} + \tau_{F,t} = g_F + \frac{M_{F,t-1}}{P_{F,t}} + R_{F,t-1} \frac{B_{F,t-1}}{P_{F,t}}. \tag{2.52}$$

The monetary authority sets the nominal interest rate R_F following the interest rate rule:

$$R_{F,t} = \alpha_{F0} + \alpha_F \pi_{F,t} + z_{F,t}, \tag{2.53}$$

where α_{0F} and α_F are policy coefficients. z_F denotes a monetary policy shock, specified

[9] All variables and parameters associated with the FTPL model are labeled with a F.

as

$$z_{F,t} = \rho_{F,1} z_{F,t-1} + \epsilon_{F1,t} \tag{2.54}$$
$$\epsilon_{F1,t} \sim N(0, \sigma_{F1}). \tag{2.55}$$

The fiscal authority sets taxes according to:

$$\tau_{F,t} = \gamma_{F0} + \gamma_F b_{F,t-1} + \psi_{F,t}, \tag{2.56}$$

where again γ_{F0} and γ denote policy coefficients. The innovation in fiscal policy has the following characteristics:

$$\psi_{F,t} = \rho_F 2\psi_{F,t-1} + \epsilon_{F2,t} \tag{2.57}$$
$$\epsilon_{F2,t} \sim N(0, \sigma_{F2}). \tag{2.58}$$

The model can be linearized and summarized by two equations[10]:

$$\tilde{\pi}_{F,t+1} = \beta_F \alpha_F \tilde{\pi}_{F,t} + \beta_F z_{F,t} \tag{2.59}$$
$$\tilde{b}_{F,t} + \varphi_{F1}\tilde{\pi}_{F,t} + \varphi_{F3} z_{F,t} + \psi_{F,t} = (\beta_F^{-1} - \gamma_F)\tilde{b}_{F,t-1} - \varphi_{F4} z_{F,t-1} - \varphi_{F2}\tilde{\pi}_{F,t-1}. \tag{2.60}$$

2.5.2 Dynamics of the FTPL model

The dynamics of the system depend on whether fiscal and monetary policy are active or passive, i.e. they depend on the policy parameters α_F and γ_F only. Different policy regimes emerge for:

- $|\alpha_F \beta_F| > 1$ and $|\beta_F^{-1} - \gamma_F| < 1$ for active monetary (AM) and passive fiscal (PF) policy. This will be referred to as regime I.
- $|\alpha_F \beta_F| < 1$ and $|\beta_F^{-1} - \gamma_F| > 1$ for active fiscal (AF) and passive monetary (PM) policy. This will be referred to as regime II.
- AM/AF and PF/PM. These cases are not considered here.

Both policy regimes imply different signs of the impulse response function for inflation and real debt. In regime I a monetary policy shock (an unanticipated increase in the nominal interest rate) will lead to a negative response of inflation and a positive response of real debt. A fiscal policy shock (an unanticipated increase in taxes) will

[10]See appendix A.2 for a derivation.

have no effect on inflation and decrease the real debt. In regime II, a monetary policy shock leads to an increase in inflation and an initial decrease in real debt. A fiscal policy shock has a negative effect on both variables. Impulse response functions for each shock, regime and variable are plotted in appendix A.5 together with the corresponding distributions of α_F and γ_F .

2.5.3 Specification and Identification of the VAR

The VAR model consists of two variables, inflation π_F and real debt b_F, with no constant or time trend: $y_{F,t} = [\pi_{F,t} b_{F,t}]'$. The VAR model with one lag is given by:

$$\begin{aligned} y_{F,t} &= B y_{F,t-1} + u_{F,t} \\ E[u_{F,t} u'_{F,t}] &= \Sigma_F. \end{aligned}$$

Ordering the fiscal policy shock first and the monetary policy shock second, based on the model the following characteristics of the impulse matrix A_F have to hold:

- If regime I holds:
  - Fiscal policy shock: $A_{F,11} = 0$ $A_{F,21} < 0$.
  - Monetary policy shock: $A_{F,21} < 0$ and $A_{F,22} > 0$.
- If regime II holds:
  - Fiscal policy shock: $A_{F,11} < 0$ $A_{F,21} < 0$.
  - Monetary policy shock: $A_{F,21} > 0$ and $A_{F,22} < 0$.

Since the sign of the reaction of real debt to a monetary policy shock does not identify the shock in the case of regime II, the monetary policy shock is ordered second, implying that both variables have to fulfil the sign restriction for a fiscal policy shock first. Then the sign of the response of real debt is restricted, while the response of inflation is left open.

2.5.4 A Monte Carlo Experiment

I simulate data from the model over 200 periods with $\alpha_F = 0.5$ and$\gamma_F = -0.00001$, i.e. the case of active fiscal and passive monetary policy. I choose the prior distribution of α_F and γ_F based on estimates by Davig and Leeper (2005):

Parameter	mean(I)	standard deviation(I)	mean(II)	standard deviation(II)
α_F	1.308	0.253	0.522	0.175
γ_F	0.0136	0.012	-0.0094	0.013

Table 2.1: Prior distribution for parameters of the FTPL model

The prior distribution is plotted in figure A.1. The model fulfills the requirements for investigating the question of how inflation reacts after a monetary policy shock: depending on the parameterization it allows for qualitatively different reactions of inflation, and the DSGE model incorporates all other shocks necessary, here the fiscal policy shock, to distinguish the shock of interest. The corresponding impulse responses for each regime are plotted in appendix A.5: Figures A.2 and A.3 provide Bayesian impulse response plots for draws from the prior distribution of regime I and figures A.4 and A.5 for draws from the prior distribution of regime II.

The sampling algorithm is specified by setting $d = 200$ to approximate the candidate distribution. Afterwards, $I = 50$ and 200 draws are taken at each iteration. Since the data are simulated from regime II, the outcome expected is the distribution of regime II, with the corresponding impulse responses of inflation and real debt for a fiscal policy shock and real debt for a monetary policy shock. Furthermore, inflation should rise in response to a monetary policy shock.

As figure A.6 indicates, this is indeed the case, even though I initialize the chain with a wrong guess. The posterior distribution of α_F and γ_F stems from regime II only. Figure A.7 shows the response to a fiscal policy shock and figure A.8 the response to a monetary policy shock. Inflation is indeed increasing.

2.6 Example 2: Application to the data

In this section I take the methodology to the data. Since the FTPL model is too stylized I introduce another very simple DSGE model: the deep habit model. This model was employed by Ravn et al. (2008) to answer the question whether a simple model can account for the so-called price puzzle, i.e. the increase of inflation after a contractionary monetary policy shock. This DSGE model is able to generate a positive as well as a negative response of inflation after a monetary policy shock.

I flip the question and explore whether prices increase or decrease after a monetary policy shock using the methodology set out in this chapter. I identify the monetary policy shock by employing sign restrictions from the DSGE model. The response of inflation will be left open.

In the remaining part of the section, first the DSGE model is set up and its dynamics are described. Finally, the DSGE model and the VAR model are estimated jointly.[11]

2.6.1 Deep habits model

The DSGE model consists of households, firms and a monetary authority. In the following, these parts of the DSGE model are characterized, the equilibrium is defined and the impulse response functions are analyzed. All variables and parameters associated with the deep habits model are labeled with an H.

2.6.1.1 Households

There is a continuum of households indexed by $j \in [0,1]$, which are all identical and infinitively lived. Household j's preferences are given by:

$$U_0^j = E_0 \sum_{t=0}^{\infty} \beta_H^t \left[\frac{1}{1-\sigma_H} x_{H,t}^j - \frac{\gamma_H}{1+\kappa_H} (h_{H,t}^j)^{1+\kappa_H} \right] \tag{2.61}$$

$$x_{H,t}^j = \left[\int_0^1 (c_{Hi,t}^j - \theta_H^d c_{Hi,t-1})^{1-\frac{1}{\eta_H}} di \right]^{\frac{1}{1-1/\eta_H}} \tag{2.62}$$

$$c_{Hi,t} = \int_0^1 c_{Hi,t}^j dj \tag{2.63}$$

where β_H denotes the discount factor, κ_H the inverse of the Frisch elasticity of labor supply, γ_H a preference weight on households j's labor supply and $x_{H,t}^j$ denotes the consumption basket. Equation (2.62) defines the deep habit: consumption of variety i is related to the past aggregate of this variety. Deep habits therefore imply that the level of marginal utility of individual goods varies. The aggregate is assumed to be exogenously given. The parameter θ_H^d measures the importance of the habit.

Demand for c_{it}^j is given by:

$$c_{Hi,t}^j = \left(\frac{P_{Hi,t}}{P_{H,t}} x_{H,t}^j + \theta_H^d c_{Hi,t-1} \right), \tag{2.64}$$

where $P_{H,t}$ denotes an aggregate price index:

$$P_{H,t} = \left[\int_0^1 P_{Hi,t}^{1-\eta_H} di \right]^{1/(1-\eta_H)}. \tag{2.65}$$

Households act as monopolistically competitive labor unions in the labor market

[11]The (uncommented) matlab codes are available upon request and will (hopefully) be made available commented on my webpage soon.

earning the wage rate W_H^j. They face costs of changing wages ζ_{Hw}, which are quadratic in the deviation of nominal wage growth from an index factor $\tilde{\pi}_{Hw,t}$. The index factor evolves according to:

$$\tilde{\pi}_{Hw,t} = \vartheta_{Hw}\pi^{\star}_{Hw} + (1-\nu_{Hw})\pi_{Hw,t-1}, \tag{2.66}$$

where ν_{Hw} measures the extent of wage indexation.

Households own firms and receive dividends $D_{H,t}^j$, and furthermore have access to a nominal risk-free bond B_H yielding the gross nominal interest rate R_H.

They maximize utility with respect to the budget constraint:

$$\begin{aligned} P_{H,t}x_{H,t}^j + \theta_H^d \int_0^1 P_{Hi,t}c_{Hi,t-1}di + B_{H,t}^j \qquad (2.67) \\ = R_{H,t-1}B_{H,t-1}^j + W_{H,t}^j h_{H,t}^j + D_{H,t}^j - P_{H,t}\frac{\zeta_{Hw}}{2}\left(\frac{W_{H,t}^j}{W_{H,t-1}^j} - \tilde{\pi}_{Hw,t}\right). \end{aligned}$$

2.6.1.2 Firms

Firms are monopolistically competitive. Firm i produces output using the following production function:

$$y_{Hi,t} = h_{Hi,t}. \tag{2.68}$$

Labor input is defined as:

$$h_{Hi,t} = \left(\int_0^1 (h_{Hi,t}^j)^{1-1/\psi_H} dj\right)^{1/(1-1/\phi_H)}. \tag{2.69}$$

Given the price $W_{H,t}^j$ for $h_{H,t}^j$, the labor demand function is given by:

$$h_{Hi,t}^j = \left(\frac{W_{H,t}^j}{W_{H,t}}\right)^{-\psi_H} h_{Hi,t}, \tag{2.70}$$

where the aggregate wage rate $W_{H,t}$ is defined as:

$$W_{H,t} = \left[\int_0^1 W_{Hj,t}^{1-\psi_H} dj\right]^{1/(1-\psi_H)}. \tag{2.71}$$

Aggregating (2.70) yields the demand for household j's labor:

$$h^j_{H,t} = \left(\frac{W^j_{H,t}}{W_{H,t}}\right)^{-\psi_H} h_{H,t}. \tag{2.72}$$

Aggregating across consumers, the demand function for firm i's product is given by:

$$\begin{aligned} c_{Hi,t} &= \left(\frac{P_{Hi,t}}{P_{H,t}}\right)^{-\eta_H} x_{H,t} + \theta^d_H \theta^d_H c_{Hi,t-1} \qquad (2.73)\\ c_{Hi,t} &= \int_0^1 c^j_{Hi,t} dj \\ x_{H,t} &= \int_0^1 x^j_{H,t} dj. \end{aligned}$$

From (2.73) the main mechanism becomes apparent: firms have an incentive to lower prices today if they expect future demand to be high relative to current demand. Additionally, the firm increases its weight on the price elastic term and therefore its price elasticity of demand.

Firms face quadratic adjustment costs ζ_{Hp} when changing nominal prices. Firms' profits are discounted by the following discount factor:

$$q_{H,t} = \beta^t_H \frac{x^{-\sigma_H}_{H,t}}{P_{H,t}}.$$

The maximization problem of firm i thus reads:

$$\underset{P_{Hi,t}}{m} ax E_0 = \sum_{t=0}^{\infty} q_{H,t} D_{Hi,t} \tag{2.74}$$

$$D_{Hi,t} = P_{Hi,t} c_{Hi,t} - W_{H,t} h_{Hi,t} - \frac{\zeta_{Hp}}{2} P_{H,t} \left(\frac{P_{Hi,t}}{P_{Hi,t-1}} - \tilde{\pi}_{H,t}\right)^2.$$

Nominal prices are indexed by $\tilde{\pi}_{H,t}$, which evolves according to:

$$\tilde{\pi}_{H,t} = \nu_{Hp} \pi^\star_H + (1 - \nu_{Hp}) \pi_{H,t-1}. \tag{2.75}$$

2.6.1.3 Monetary Policy and market clearing

Monetary policy aims at stabilizing deviations in inflation and output from their steady state values $\pi^\star_H$ and $y^\star_H$. It sets the policy coefficients ρ_{Hr}, $\alpha_{H,\pi}$ and $\alpha_{H,y}$ according to the simple interest rate rule:

$$R_{H,t} = R_H^\star + \rho_{H,r}(R_{H,t-1} - R_H^\star) + (1-\rho_{Hr}) \left[\alpha_{H,\pi}(\pi_{H,t} - \pi_H^\star) + \alpha_{Hy} \left(\frac{y_{H,t} - y_H^\star}{y_H^\star} \right) + \epsilon_{H,t} \right]. \tag{2.76}$$

$\epsilon_{H,t}$ denotes the monetary policy shock: $\epsilon_{H,t} \sim \mathcal{N}(0, \sigma_{HR})$.

Market clearing implies:

$$h^j_{H,t} = \int_0^1 h^j_{Hi,t} di \tag{2.77}$$

$$h_{Hi,t} = \int_0^1 h^j_{Hi,t} dj \tag{2.78}$$

as well as:

$$c_{H,t} = y_{H,t}. \tag{2.79}$$

The aggregate resource constraint is given by:

$$y_{H,t} = h_{H,t}. \tag{2.80}$$

2.6.1.4 Equilibrium definition

I follow Ravn et al. (2008) by concentrating on the symmetric equilibrium in which all consumers make the same choice over consumption, set the same wage and all firms set the same prices.

A recursive equilibrium is then defined as follows:

Definition 1 *Given initial values $P_{H,0} > 0$ and $W_{H,0}$, the recursive laws of motion for price and wage indexation* (2.75) *and* (2.66) *and a monetary policy, a rational expectations equilibrium (REE) for $R_{H,t} \geq 1$, is a set of sequences $\{y_{H,t}, c_{H,t}, h_{H,t}, x_{H,t}, w_{H,t}, P_{H,t}, R_{H,t}\}_{t=t_0}^{\infty}$*

(i) that solve the firms' problem (2.74) *with s.t.* (2.73)*,*

(ii) that maximize households' utility (2.61) *s.t.* (2.72), (2.67) *and a No-Ponzi-scheme condition,*

(iii) that clear the goods market (2.79) *and labor market, i.e.* (2.78) *and* (2.77) *hold,*

(iv) and that satisfy the aggregate resource constraint (2.80)*.*

The DSGE model is loglinearized around its steady state. An overview of the steady state and the loglinearized equations are given in Appendix A.3.1 and A.3.2 respectively.

2.6.1.5 Prior distribution of the parameters and impulse response functions

I estimate only those structural parameters crucial for the response of inflation[12]. For those parameters, prior distributions are specified which allow for a wide range of impulse response functions of the deep habits model. The parameters not estimated are calibrated as in Ravn et al. (2008). An overview can be found in appendix A.3.3.

2.6.2 Estimation

In figure A.9 the impulse response functions of the DSGE model when drawing from the prior distribution are plotted.[13] The signs of all the impulse response functions except the response of interest (inflation) are constant, i.e. for every draw from the parameter distribution of the DSGE model consumption, real wages and output will be decreasing while the interest rate increases. In order to distinguish the characterization of the shock from other shocks, I compare the combination of signs with combinations implied by other common shocks. These shocks are taken from Smets and Wouters (2003). The sign restriction of the monetary policy shock implied by the deep habits model are different from the signs of common shocks except for the price markup shock in Smets and Wouters (2003). Even though it is the shock exhibiting the smallest variance, I further include adjusted reserves as well as the price index of crude materials into the VAR model to distinguish the estimated shock (following Mountford and Uhlig (2005)). While the former is restricted to react negativly, the latter is left unrestricted. Since both variables have no counterparts in the DSGE model, they are not matched. Overall, the VAR model consists of 7 variables: real GDP, real personal consumption, real wages, interest rates, adjusted reserves, the GDP deflator and the price index of crude materials. A complete description of the time series is given in Appendix A.4. The prior distribution of the VAR model is specified as a flat prior.

Before the DSGE model is estimated, I perform a Monte Carlo experiment to ensure the validity of the methodology, the identification and the specification of the sampling algorithm. The candidate distribution for the vector of deep parameters will be the prior distribution. In the Monte Carlo experiment I set $I = 20$ and draw $n = 200$ times at each iteration. First only a subvector of the parameters of the DSGE model consisting of θ_d, η, ζ_w, and ζ_p is estimated. The results are displayed in table A.1 (columns 6 and 7) of appendix A.3.3 and show that all parameters are estimated

[12]Ravn et al. (2008) also only estimate a subset of the structural parameters.

[13]All figures are provided in appendix A.5.

very precisely around their true values (column 5). This is a very encouraging result, especially since the prior distribution is not centered around the true value.

Adding more parameters to the vector of estimated parameters has two effects. This is demonstrated by supplementing the vector of structural parameters with the coefficients of the Taylor rule ($\rho_{Hr}, \alpha_{H\pi}, \alpha_{Hy}$) and the inflation indexation parameter ν_{Hp}. On the one hand, this increases the flexibility of the DSGE model and therefore increases the ability to fit the impulse response functions of the data. Figure A.10 provides plots of the impulse response function of the DSGE model and the VAR model. Both coincide and, more importantly, the 'true' impulse response function for inflation, i.e. the impulse response function for the parameter vector at which the DSGE model is simulated, is estimated. On the other hand, as shown in table A.1 columns 8 and 9, the precision of the estimation is slightly blurred.

Given the encouraging results, I take the methodology to the data. At every iteration I take $n = 200$ draws, the number of iterations is set to $I = 20$. Table A.1 column 10 and 11 report the mean and the standard deviation of the posterior distribution respectively. The estimation results for the posterior mean of some of the parameters of the DSGE model are very similar to those obtained by Ravn et al. (2008)[14]: $\eta_H = 2.47$ (2.48), $\zeta_{Hp} = 14.89$ (14.47), $\zeta_{Hp} = 42.50$ (40.89),$\alpha_{Hr} = 0.01(0.04)$. I find slightly different estimates for the deep habit parameter $\theta^d_H = 0.72$ (0.85), the inflation indexation parameter $\nu_{Hp} = 0.1$ (0), and the policy coefficients $\rho_{Hr} = 0.81$ (0.74) and $\alpha_{H\pi} = 1.56, (1.26)$. Figure A.12 displays the impulse response functions: since the parameters of the DSGE model are estimated similarly, the response of inflation is positive and significant for 68% probability bands.[15] However, while the graph indicates a positive response for the mean response of the VAR model, the uncertainty bands give rise to the conclusion that a negative response of inflation to a monetary shock is as likely as positive one.

2.7 Conclusion

This chapter has laid out a methodology for identifying the structural shocks of a Vector Autoregression (VAR) model while at the same time estimating a Dynamic Stochastic General Equilibrium (DSGE) model that is not assumed to replicate the data-generating process. To this end it has presented a framework for jointly estimating the parameters of a VAR model and a DSGE model.

[14]For comparison I report their findings in brackets after my estimates.

[15]It is not significant for 100% probability bands.

The VAR model is identified based on restrictions from the DSGE model, i.e identification relies on restrictions explicitly derived from theory. This ensures consistency of the identification of the VAR model with the implied structural form of the DSGE model. Restrictions are formulated as sign restrictions. Thus, the DSGE model serves as a way to summarize the ideas economists have about the economy. Ideally, it incorporates the assumptions the researcher wants to discriminate between, but in any case it should be as agnostic as possible about the response of the variables of interest to the shock of interest.

The DSGE model is estimated by matching the impulse response functions of the VAR and of the DSGE, i.e. their implied dynamics. Therefore, it need not be a representation of the data-generating process. While the shock of interest has to be included, as well as other shocks necessary to distinguish it, the DSGE model need not be fully stochastically specified.

The methodology has been first illustrated by means of a Monte Carlo experiment and has been applied to the data afterwards. In the Monte Carlo experiment, artificial data has been simulated from a simple fiscal theory of the price level model in which fiscal policy is active and monetary policy passive. The sign of the response of inflation to a monetary policy shock has been investigated. Depending on the policy regime, i.e. the reaction coefficients of the policy rules, the response can either be negative or positive. The prior distributions of the policy parameters have been chosen such as to ensure that both regimes and therefore both responses are equally likely. The estimated impulse response function of the VAR model as well as the posterior distribution of the parameter of the DSGE model indicate that the methodology works correctly: the response of inflation shows the 'true' sign and the posterior distribution of the parameter of the DSGE model consists solely of policy coefficients from active fiscal and passive monetary policy.

Finally, the methodology has been used to estimate the response of inflation to a monetary policy shock. As a DSGE model, the deep habits model laid out by Ravn et al. (2008) has been employed. The posterior estimates of the parameters of the DSGE model are similar or only slightly different from those obtained by the authors. Correspondingly, I find a positive and on a 68% level significant response of inflation to a monetary policy shock. However, while the mean of the impulse response function of the VAR model is positive, the uncertainty bands indicate that a negative response of inflation to a monetary policy is as likely as a positive one.

Chapter 3

Pre-announcement and Timing – The Effects of a Government Expenditure Shock

This chapter investigates the effect of a government expenditure shock on consumption and real wages. I identify the shock by exploiting its pre-announced nature, i.e. different signs of the responses in investment, hours worked and output during the announcement and after the realization of the shock. Since pre-announcement leads to a non-stationary moving average representation, I estimate and identify a VMA model. The identifying restrictions are derived from a DSGE model, which is estimated by matching the impulse response functions of the VMA model. Private consumption is found to respond negatively during the announcement period and positively after the realization. The reaction of real wages is significantly positive on impact, decreases during the announcement horizon, and is again significantly positive for two quarters after the realization.

3.1 Introduction

What are the effects of an innovation in government expenditure on the economy? It still remains disputed whether private consumption and real wages rise or fall in response. While economic theory supports both outcomes, empiricists have not yet been able to discriminate between different explanations, as the literature is divided over the appropriate methodological approach. On the one hand, the narrative approach identifying changes in government spending by military buildup dates in the US, predicts a decrease in consumption and real wages. On the other hand, the structural Vector Autoregression (SVAR) approach identifying the shocks either with timing or sign restriction, predicts the opposite.

In this chapter I investigate the effects of a government expenditure shock by employing an SVAR model approach and identifying it via sign restriction as in Mountford and Uhlig (2005) and Pappa (2005). Moreover, I explicitly model pre-announcement of the fiscal policy shock and its consequences: the behavior of investment, GDP and hours worked differs during the announcement period and after the realization of the shock[1]. By doing so, I first avoid both criticism of the narrative approach - the potential small sample problems, be it with one shock or a combination of many - and also the ongoing discussion whether 'abnormal' fiscal policy episodes lead to 'abnormal' or 'normal' behavior of the economy.

Second, the approach set out in the paper also sidesteps the potential pitfalls of the SVAR approach. In a recent paper, Ramey (2008) pointed out that the different results obtained by using this approach are simply due to a faulty timing assumption in the SVAR literature which neglects or misses that changes in the government budget are often pre-announced or known to the public beforehand. One recent example is The American Recovery and Reinvestment Act of 2009, which announces government expenditures from 1009 to 2012 with most spending taken place in 2010 and 2011. As I will demonstrate in this paper, identifying the government expenditure shock by exploiting the pre-announcement effects explicitly, i.e. considering qualitative differences in the response of investment during the pre-announcement period and after the realization of the shock, overcomes the critique by Ramey.

When pursuing this approach another issue arises: the moving average representation of the data generated by a pre-announced policy is potentially non-stable so that it cannot be approximated by a VAR model[2]. Non-stability of the moving average representation has two consequences: First, the information set of the agent in the economy is larger than the information set of the econometrician: the agent has news about future, pre-announced changes, which are not contained in the information set observed by the econometrician. This news is discounted by the agent in a different way than by the econometrician. The econometrician estimates innovations as a discounted sum of past news, where older news receives less weight than more recent. But, for an agent yesterday's news on pre-announced policy can be more relevant than today's.

[1]The existence of those differences were found by House and Shapiro (2006) and Trabandt and Uhlig (2006) among others.

[2]As pointed out first by Hansen and Sargent (1991) and more recently by Leeper et al. (2008).

One challenge is therefore to align the information set of the econometrician with that of the agent. In this paper I achieve this by using additional information taken from the DSGE model. A second difficulty arises when attempting to estimate the non-stable moving average process. One possibility, as laid out by Lippi and Reichlin (1993) and Lippi and Reichlin (1994) is to flip the unstable roots of the moving average process into a stable region via Blaschke factors. Another way to proceed is to estimate the moving average process with a Kalman Filter[3]. In this paper I therefore estimate the Vector moving average representation.

Since there exists no common knowledge about the behavior of the macroeconomic variables, I formally derive the sign restrictions from a dynamic stochastic general equilibrium (DSGE) model. As DSGE model, I employ a model laid out by Galí, López-Salido, and Vallés (2007). It is well suited to resolving the debate because it addresses the arguments why the classic results should hold as well as the typical arguments why the classic results fail: households which cannot smooth consumption, imperfect labor markets and a certain degree of price stickiness. Depending on its parametrization, for example the fraction of rule-of-thumb consumers and the degree of price stickiness, the model features several classic or Keynesian characteristics. In the limit, i.e. with no rule-of-thumb consumers and firms allowed to reset prices each period, it boils down to a neoclassical model. To what extend those components influence the variables of interest and the resulting sign restrictions for the VMA model depends on a small number of structural parameters. The parameters of the DSGE model are estimated jointly.

The results for a three-quarter pre-announced increase in government expenditures show strong qualitative differences during the announcement period and after the realization of the shock: output and hours worked respond negatively during the announcement period and positively afterwards; investment responds negatively one additional quarter before responding positively. Private consumption mimics this behavior and shows a stable, slightly negative response during the announcement period followed by a significant positive response after the realization of the shock. Real wages react significantly positively on impact, decrease (and even become negative) during the announcement horizon and react significantly positively for two quarters after the realization.

The chapter is organized in the following way: the next section summarizes the related literature. In section 3.3 I lay out the econometric strategy. The DSGE model

[3]This is shown by Eric Leeper on slides available on his webpage. Appendix B.3 summarizes his insights to keep the paper self-explanatory.

used to derive the sign restriction is described in section 3.4. The results are summarized in section 3.5. Section 3.6 concludes.

3.2 Related Literature

In this section I first shortly review the theoretical work on the effects of a government expenditure shock before I discuss the existing empirical approaches and findings in more detail.

The evolution of the literature on the effects of a government expenditure shock can be summarized in the following way: starting out from a neoclassic growth model, which is step by step extended with market imperfections and nominal rigidities, an ultra New Keynesian model (as Ramey called it) evolved. The first neoclassic attempts to study the effects of fiscal policy date back to Hall (1980) and Barro (Barro (1981) and Barro (1987)). Building upon this work, Aiyagari, Christiano, and Eichenbaum (1990) and Baxter and King (1993) expand a neoclassical growth model of a government sector. In these models, an increase in government expenditure creates a negative wealth effect for the household, which will reduce consumption and increase labor supply. The increased labor supply induces real wages to decrease and interest rates to increase.

Rotemberg and Woodford (1992) and Devereux, Head, and Lapham (1996) introduced market imperfections, increasing returns to scale as well as monopolistic and oligopolistic competition respectively into the neoclassic growth model. In their models, a government spending shock, or a general demand shock, increases demand for goods, thereby labor demand and thus real wages. In a recent paper, Galí et al. (2007) extend the New Keynesian model with rule-of-thumb consumers, who neither borrow nor save, only consume the disposable income each period. Since those households do not feel intertemporally poorer, they do not decrease consumption as a response to a positive government expenditure shock. Confronted with this large number of competing models, empiricists have tried to discriminate between them by investigating the response of real wages and private consumption after a change in government spending.

Findings in existing empirical studies are twofold depending on the identification scheme of the government expenditure shock. Ramey and Shapiro (1998) use a narrative approach to identify the VAR model. They interpret times of large military buildups in the US, the Korean war, the Vietnam war and the Carter-Reagan buildup, as sudden and unforseen increases in government expenditure. The resulting reactions of macroeconomic variables to these events are thus interpreted as deviations from normal behavior. They find that output and hours rise, while consumption and real wages fall. Burnside et al. (2004) employ a similar methodology to estimate the impulse responses of macroeconomic variables to a government expenditure shock and compare them to impulse responses implied by a standard neoclassic model. The results indicate that hours worked rise and investment briefly increases, while real wages and consumption decrease. Thus they conclude that the standard neoclassic model can account reasonably well for the effects of fiscal policy shocks. A similar conclusion is drawn by Edelberg et al. (1999), who modify a neoclassic growth model distinguishing two types of capital, nonresidential and residential.

A structural VAR approach is chosen by Blanchard and Perotti (2002) to identify a government expenditure shock. They require fiscal policy variables not to respond immediately to other innovations in the economy, i.e. they employ the recursiveness assumption. Their findings corroborate the results of Ramey and Shapiro (1998) concerning output and hours worked, but contradict their findings for consumption and real wages. Mountford and Uhlig (2005) also use a structural VAR, but do not consider any timing restriction. Instead, they employ sign restrictions to restrict the responses of fiscal variables, while the responses of other macroeconomic variables are left open. Besides the different methodology, they additionally allow for a pre-announcement of fiscal policy shocks. Indeed, as has been widely acknowledged and mentioned, most fiscal policy shocks are pre-announced. Their findings, however, confirm the results of Blanchard and Perotti (2002) except for consumption, which only shows a weak positive response.

The debate about the empirical evidence was reopened by Ramey (2008)[4]. Her paper takes up two issues. First, she stresses the importance of the composition of government expenditures. The dataset used by Blanchard and Perotti (2002) includes government consumption as well as government investment expenditure. An increase in the latter can be productive and potentially complement private consumption and investment so that it might lead to a positive response in those variables. For these reasons Ramey advocates using defense spending as a proxy for government expenditures in the VAR. Second, she states that the findings of the studies differ due to pre-announcement effects, implying that Blanchard and Perotti (2002) employ faulty timing to identify the fiscal policy shock. In her paper, a neoclassic DSGE model including a pre-announced government expenditure shock is set up and used to simulate artificial data. It is then demonstrated that, if the pre-announcement of the shock is taken into account, a negative response in consumption is estimated. If not, consumption appears to react positively, a clearly misleading result.

[4]The first version dates back to 2006.

In his summary and discussion of the recent literature, Perotti (2007) acknowledges the concerns with respect to the structural VAR methodology. As a possible way of overcoming its weaknesses he suggests employing annual data and distinguishing between shocks to defense spending and to civilian government spending. However, using annual data, the recursive assumption that the fiscal sector does not react contemporarily to the state of the economy might not hold anymore. But, as Perotti mentions, the narrative approach has considerable weaknesses of its own: First, it suffers from a small sample size, second, it is not entirely clear whether the whole change in government expenditure was announced at once or whether it was a combination of small changes, i.e. whether there were numerous revisions of the military budget, occurring one after the other, causing private consumption to respond multiple times.

In Ravn et al. (2007), the authors dismiss Ramey's critique of the usage of structural VAR models. They point out that shocks are by assumption orthogonal to the information set and consequently identify a structural VAR as in Blanchard and Perotti (2002).

Pre-announced fiscal policy is considered in Mertens and Ravn (2009) and Tenhofen and Wolff (2007). The former authors address the issue of non-invertibility of the moving average representation by using a Vector Error Correction Model. The government expenditure shock is then identified by a combination of long run and zero restrictions. The latter authors augment the original VAR model by Blanchard and Perotti (2002) with expectations and consider a one period pre-announcement only. This paper differs in two dimensions: first by using sign restriction to identify the government expenditure, putting less structure and fewer restrictions on the VAR model, and second by estimating a VMA model to circumvent the issue of non-invertibility.

Besides the problem of non-invertibility of the moving average representation, Chung and Leeper (2007) discuss the importance of the intertemporal government budget constraint for a structural VAR analysis. In order to estimate reduced form shocks that can be mapped into structural innovations, government debt and private investment should be included in a VAR model.

Another important issue to take into account, as pointed out by Ramey (2008), is the composition of government expenditure and what part is used in the estimation. Abstracting from government transfers, government expenditure is defined as the sum of government investment expenditure and government consumption expenditure. Both types have different implications for the variables of interest. As described in Turnovsky and Fisher (1995), an increase in government investment expenditure increases productivity and therefore private consumption and real wages. Including government investment expenditure in the analysis would therefore favor a Keynesian outcome, i.e. an increase in both variables of interest and an adulteration of the analysis. Furthermore, government consumption expenditure is quantitatively much more relevant (since the 1970s it has been about five times as large as government investment expenditure). I will therefore employ government consumption expenditure when estimating the effects of an innovation in government expenditure.

3.3 Econometric Strategy

Before describing the estimation methodology in detail, I will discuss one note of caution when estimating fiscal policy set out by Leeper et al. (2008). They discuss the problem that private agents have foresight about future fiscal policy, which the investigating econometrician does not have, i.e. the information set of the private agent is larger than the information set of the econometrician. This leads to the effect, that agents discount news differently compared to the econometrician. More precisely, the econometrician estimates innovations as a discounted sum of old news where former news receives less weight. For the agent former news has a larger effect on today's variables due to pre-announcement. Secondly, they show that whenever foresight about future fiscal variables is present the resulting moving average representation of the data exhibits roots inside the unit circle. Consequently, the moving average process is not invertible and cannot be approximated by a VAR model. Furthermore, estimation of a moving average process with this kind of long memory property proves to be very difficult.

There exist two possibilities to address the latter issue: flipping the roots of the moving average process outside the unit circle by either applying a Blaschke-factor or a Kalman Filter.[5] I will employ the latter and thus obtain correct estimates of the reduced form moving average coefficients[6]. In order to recover the information set of the private agent I will apply sign restrictions.

These restrictions are derived from impulse response functions of the DSGE model. Two issues become crucial: the choice of the DSGE model and its parametrization. As DSGE model I employ a model laid out by Galí et al. (2007). It is well suited to navigating through the debate because it addresses the arguments why the classic results should hold as well as the typical arguments why the classic results fail: households

[5]Leeper et al. (2008) discuss both possibilities extensively. Blaschke factors are used by Mertens and Ravn (2009).

[6]More algebraic details are given in appendix B.3.

which cannot smooth consumption, imperfect labor markets and a certain degree of price stickiness. Depending on its parametrization, for example the fraction of rule-of-thumb consumers and the degree of price stickiness, the model features several classic or Keynesian characteristics. In the limit, i.e. with no rule-of-thumb consumers and firms allowed to reset prices each period, it boils down to a neoclassical model. Therefore, as Figure B.3 indicates, this DSGE model allows for positive as well as negative responses in consumption and real wages. To what extend these components influence the variables of interest and the resulting sign restrictions for the VMA model depends on a small number of structural parameters, which have to be estimated.

The estimation methodology for the structural parameters of the DSGE model is chosen on the following grounds: due to advances in computational power there are various ways to estimate a DSGE model nowadays. However, procedures and results differ substantially depending on the econometric interpretation of the DSGE model. Geweke (1999a) distinguishes between a strong and a weak interpretation. The former requires the DSGE model to provide a full description of the data generating process. Formally, it has to exhibit as many structural shocks as the observable variables to be explained. The DSGE model described above is clearly not intended to be a proper representation of the data generating process[7]. One estimation strategy therefore is to extend the DSGE model with additional structural shocks and several additional features and frictions such as for example capacity utilization and a habit in consumption. This is no such thing as a free lunch. This comes at the cost of a diluted analysis. The result cannot be traced back to certain DSGE model components such as for example the share of rule-of-thumb households and the degree of price stickiness, whose effects are to be investigated. Attempting to apply the strong econometric interpretation to DSGE model without extensions results in biased estimation results.[8] I therefore follow the weak econometric interpretation of the DSGE model and do not assume that it is a proper representation of the data generating process. Instead, the DSGE model is estimated by matching the impulse response functions of the DSGE model and a VMA model to a government expenditure shock by the methodology laid out in detail in chapter 2.

Thus, on the one hand the parameters of the DSGE model are estimated by matching the corresponding impulse response functions of the VMA model. On the other hand, the structural impulse response functions of the VMA model are identified by applying sign restrictions which are derived from the DSGE model. Both distributions

[7]As is pointed out by Galí et al. (2007).

[8]In that case it is not possible to identify all structural parameters.

are therefore conditional distributions: they depend on a realization of the impulse response function of the VMA model and on restrictions from the DSGE model, i.e. a realization of a vector of structural parameters of the DSGE model and a realization of the coefficients of the VMA model, respectively. To take those dependencies into account it is necessary to consider both, the impulse response functions of the VMA model and of the DSGE model as stochastic and to characterize their joint distribution.

There is one aspect, in which the application of the methodology differs from chapter 2. In chapter 2 the VMA model was approximated by a VAR model. In this chapter I estimate the VMA model. This yields a slightly modified estimation of the time series model as well as a modified sampling algorithm. Both are described in detail in Appendix B.2.2 and B.2.3.

3.4 The DSGE model

The DSGE model employed here was originally laid out by Galí et al. (2007). I relax the assumption of no initial debt in order to contrast the DSGE model with the data. The DSGE model consists of two types of households, households with access to government bond and capital markets and households without. Goods are produced by perfectly competitive firms, using goods produced by intermediate firms as inputs. Intermediate good firms are monopolistic competitors and subject to a Calvo pricing mechanism, i.e. with a certain probability they receive a signal allowing them to reset their price which they are not allowed to change otherwise. Intermediate good firms have access to a production technology combining capital and labor. Labor is supplied by labor unions.

The government consists of a monetary authority setting the nominal interest rate and a fiscal authority issuing bonds and raising taxes. Government expenditure is modeled as an exogenous process.

The DSGE model exhibits a continuum of infinitively lived households indexed by $i \in [0, 1]$. There exist two types of households, households optimizing intertemporally and households which are not allowed to save. The latter are called rule-of-thumb households. More precisely, they are not allowed to participate in the capital and bonds goods market, an assumption made to reflect the fact that some households have indeed only limited access to the credit market. The share of rule-of-thumb households in the economy is given by λ.

Intertemporally optimizing households maximize utility depending on consumption

c^o and labor n^o:

$$U = E_0 \sum_{t=0}^{\infty} \beta^t \left(log(c_t^o) - \frac{n_t^{o\,1+\nu}}{1+\nu} \right), \tag{3.1}$$

where β is the discount factor and ν a parameter measuring the disutility of labor. The household buys government bonds B yielding the return R, invest i^o units in capital goods k^o yielding return R^k and receives dividends d from the ownership of firms. With real wages denoted by w and lump sum taxes by t^o the budget constraint is given by:

$$c_t^o + i_t^o + \frac{B_t}{p_t R_t^b} = w_t n_t^o + R_t^k k_{t-1}^o + \frac{B_{t-1}}{p_t} + t_t^o + d_t. \tag{3.2}$$

Capital accumulates according to:

$$k_t^o = (1-\delta) k_{t-1}^o + \Phi \left(\frac{i_t^o}{k_{t-1}^o} \right) k_{t-1}^o, \tag{3.3}$$

where Φ denotes investment to capital adjustment costs. With respect to the adjustment cost function the following assumptions are made: $\Phi(\delta) = \delta$, $\Phi'(\delta) = 1$, $\Phi' > 0$, $\Phi'' \leq 0$. The elasticity of the investment to capital ratio with respect to Tobin's $q_t = \frac{1}{\Phi'\left(\frac{i_t^o}{k_{t-1}^o}\right)}$ is defined as: $\eta \equiv -\frac{1}{\Phi''(\delta)\delta}$.

Rule-of-thumb households maximize each period's utility:

$$U = log(c_t^r) - \frac{n_t^{r\,1+\nu}}{1+\nu} \tag{3.4}$$

subject to the budget constraint:

$$c_t^r = w_t n_t^r - t_t^r \tag{3.5}$$

Both types of households are assumed to consume the same amount of goods in the steady state. This can be achieved through the appropriate choices of the lump sum taxes t^r and t^o. Aggregate variables, i.e. aggregate consumption c, labor n, capital k and investment i are defined as the weighted average of the corresponding variables of the rule-of-thumb household and the intertemporally optimizing household:

$$c_t = \lambda c_t^r + (1-\lambda) c_t^o \tag{3.6}$$

$$n_t = \lambda n_t^r + (1-\lambda) n_t^o \tag{3.7}$$

$$i_t = (1-\lambda) i_t^o \tag{3.8}$$

$$k_t = (1-\lambda) k_t^o. \tag{3.9}$$

3.4.1 Firms

The economy contains two sectors. In one sector perfectly competitive firms produce the final good y using as inputs intermediate goods produced by monopolistically competitive firms.

3.4.1.1 Final good firms

Final good firms have access to the following production function:

$$y_t = \left(\int_0^1 y_t(j)^{\frac{\varpi_p - 1}{\varpi_p}} dj \right)^{\frac{\varpi_p - 1}{\varpi_p}}, \tag{3.10}$$

where $y(j)$ denotes the intermediate good produced by firm j and $\varpi_p > 1$ the elasticity of substitution between different intermediate goods. Demand for each good $y(j)$ is given by

$$y_t(j) = \left(\frac{p_t(j)}{p_t} \right)^{-\varpi_p} y_t. \tag{3.11}$$

Cost minimization under perfect competition yields

$$p_t^{1-\varpi_p} = \left(\int_0^1 p_t(j)^{1-\varpi_p} dj \right). \tag{3.12}$$

3.4.1.2 Intermediate good firms

An intermediate good firm j, $j \in [0,1]$ produces the good $y(j)$ using the production function:

$$y_t(j) = k_{t-1}(j)^{\alpha} n_t(j)^{1-\alpha}, \tag{3.13}$$

where α denotes the capital share in the production. Taking the real wage and capital as given, cost minimization implies:

$$\frac{k_t(j)}{n_t(j)} = \left(\frac{\alpha}{1-\alpha} \right) \left(\frac{w_t}{R_t^k} \right). \tag{3.14}$$

Real marginal costs MC common to all firms can be derived as

$$MC_t = \alpha^{-\alpha}(1-\alpha)^{-(1-\alpha)} R_t^{k\alpha} w_t^{1-\alpha}. \tag{3.15}$$

Intermediate good firms are subject to a Calvo pricing mechanism. Each period the firm receives a price signal with probability ϑ. The intermediate firm is allowed to set the price $p_t(j)$ and maximizes

$$E_t \sum_{k=0}^{\infty} \vartheta^k \beta^k \left(y_{t+k}(j) (\frac{p_t(j)}{p_t} - mc_{t+k}) \right) \tag{3.16}$$

subject to the demand function:

$$y_t(j) = \left(\frac{p_t(j)}{p_t} \right)^{-\varpi_p} y_t. \tag{3.17}$$

3.4.2 Government sector

The government sector consists of a monetary authority setting nominal interest rates R and a fiscal authority setting lump sum taxes t and issuing nominal government bonds B. The monetary authority follows the simple interest rate rule:

$$R_t = \phi_p i \pi_t. \tag{3.18}$$

The fiscal authority has to balance the government budget constraint given by:

$$t_t + \frac{B_t}{p_t R_t} = g_t + \frac{B_{t-1}}{p_t}. \tag{3.19}$$

Since the fiscal authority can adjust bonds and lump sum taxes, an additional fiscal policy rule is needed to determine both variables. The fiscal policy rule is of the form[9]

$$\frac{t_t}{\bar{t}} = \phi_b \frac{b_{t-1}}{\bar{b}} + \phi_g \frac{g_t}{\bar{g}}. \tag{3.20}$$

3.4.3 Labor unions

As an additional friction, an imperfect labor market is assumed. The introduction of wage-setting by unions leads to the outcome that the amount of labor supplied by the households is equal across households, i.e. $n_t^o = n_t^r$. A continuum of unions is assumed representing a certain type of labor. Let ϖ_w denote the elasticity of substitution across different types of labor. Effective labor input hired by firm j is then given by:

$$n_t(j) = \left(\int_0^1 n_t(j,i) di \right)^{\frac{\varpi_w}{\varpi_w - 1}}. \tag{3.21}$$

[9] $\bar{x}$ denote steady state values.

Assuming furthermore that the proportion of rule-of-thumb households is uniformly distributed across households and therefore across unions, a typical union (z) sets the wage in order to maximize the following expression:

$$\lambda \left[\frac{w_t(z) n_t(z)}{c_t^r(z)} - \frac{n_t(z)^{1+\nu}}{1+\nu} \right] + (1-\lambda) \left[\frac{w_t(z) n_t(z)}{c_t^o(z)} - \frac{n_t(z)^{1+\nu}}{1+\nu} \right] \tag{3.22}$$

subject to:

$$n_t(z) = \left(\frac{w_t(z)}{w_t} \right)^{-\varpi_w} n_t. \tag{3.23}$$

3.4.4 Market clearing and equilibrium

The goods market clearing condition is given by:

$$y_t = c_t + i_t + g_t. \tag{3.24}$$

An equilibrium is defined in the following way:

Definition 2 *An equilibrium is an allocation $\{c_t, n_t, i_t, k_t, b_t, y_t\}$ and a price system $\{p_t(j), p_t, w_t(z), w_t, R_t^k, R_t^b\}$ and an inflation rate π_t such that for a monetary policy R_t and fiscal policy t_t, an initial price level p_{-1}, initial values for k_{-1} and b_{-1} and given an exogenous processes for government expenditure $\{g_t\}$:*

i *for each intertemporally maximizing household an allocation $\{c_t^o, n_t^o, k_t^o, x_t^o\}$ maximizes* (3.1) *subject to the budget constraint* (3.2) *and the capital accumulation equation* (3.3)*, given prices $\{w_t, R_t^k, R_t^b\}$ and profits $\{d_t\}$;*

ii *for each rule-of-thumb household an allocation $\{c_t^r, n_t^r\}$ maximizes* (3.4) *subject to the budget constraint* (3.5)*, given prices $\{w_t\}$;*

iii *the definitions*

$$c_t = \lambda c_t^r + (1-\lambda) c_t^o \tag{3.25}$$

$$n_t = \lambda n_t^r + (1-\lambda) n_t^o \tag{3.26}$$

$$i_t = (1-\lambda) i_t^o \tag{3.27}$$

$$k_t = (1-\lambda) k_t^o \tag{3.28}$$

hold;

iv *the production allocation* $\{y_t(j), y_t\}$ *and prices* $\{p_t, p_t(j), R_t^k, w_t\}$ *solve the cost minimization problem of the final good firms and the profit maximization problem of each intermediate firm* j *subject to the demand function (3.17) and technology (3.13);*

v *for each labor union* z *the allocation* $\{n_t, n_t(z), c_t^r(z), c_t^o(z)\}$ *and prices* $\{w_t(z), w_t\}$ *maximize the pay-off function (3.22) subject to the demand for each labor union (3.23);*

vi *the government budget constraint (3.19) is fulfilled;*

vii *markets are clear.*

3.5 Results

3.5.1 Pre-announcement and timing – A Monte Carlo Study

In this section the DSGE model is calibrated to mimic the DSGE model employed by Ramey (2008). I simulate artificial data from it and estimate a structural VAR model with this dataset taking pre-announcement effects into account.

In order to mimic the DSGE model laid out by Ramey, which is a neoclassical growth model with government spending and non-distortionary taxes, the DSGE model from section 3.4 is calibrated in the following way. The fraction of rule-of-thumb households is set to 0 ($\lambda = 0$), the elasticity of the investment to capital ratio with respect to Tobin's q is set fairly high ($\eta = 10$), the probability of not optimizing prices is set very low to mirror almost flexible prices ($\vartheta = 0.05$) and the elasticity of substitution between intermediate goods is set to 1 ($\mu_p = 1$). Remaining parameters are chosen similar to Ramey, i.e. the discount factor $\beta = 0.99$, the depreciation rate $\delta = 0.025$, the capital share in production $\alpha = 0.33$ and the parameter measuring the disutility of labor $\nu = 1$. The policy parameters of the fiscal and monetary authorities are set such as to ensure uniqueness of the solution: $\phi_b = 0.3$, $\phi_g = 0.1$ and $\phi_\pi = 1.5$. The processes for actual government expenditures g_t and the forecast of government expenditure g_t^f are taken from Ramey:

$$\hat{g}_t = \hat{g}_{t-2}^f \tag{3.29}$$

$$\hat{g}_t^f = 1.4\hat{g}_{t-1}^f - 0.18\hat{g}_{t-2}^f - 0.25\hat{g}_{t-3}^f + \epsilon_t^g \tag{3.30}$$

Since I am going to estimate a VAR model with four variables, the DSGE model from which the data is simulated has to incorporate at least this number of shocks. I therefore augment the model with a preference shock to labor, a shock to total factor productivity (as Ramey does) and a monetary policy shock (i.e. a shock to equation (3.18)).

The resulting impulse response function of the DSGE model to a shock in government expenditure is plotted in Figure B.1. It is very similar to the figure in Ramey (2008). Note that investment first reacts positively to the shock and becomes negative afterwards. In order to provide evidence that making use of the pre-announced nature of a government expenditure shock, i.e. restricting variables to respond differently during the announcement of the shock and after realization ,can resolve the problem of flawed estimation approaches, I identify the government expenditure shock employing the following sign restrictions: government expenditure are assumed to be zero during the first period (pre-announcement period) and to be positive afterwards, hours worked are assumed to respond positively and investment is restricted to respond positively during the first quarter and to respond negatively afterwards.[10]

The result of the Monte Carlo Experiment is shown in Figure B.2. The response of consumption is significantly negative. This finding is robust with respect to a faulty pre-announced period (one only instead of two) and whether the government expenditure shock is ordered first or second. Given this encouraging result I will apply the methodology to the data. This time, however, the sign restrictions will not be identical across the parameter space of the DSGE model.

3.5.2 Data

I now apply the methodology to quarterly US data ranging from the first quarter of 1948 to the third of 2007. The VMA model consists of seven variables: government consumption expenditure, real GDP, private consumption, hours worked, private investment, real wages and real federal debt. The data was obtained from the internet, mostly from NIPA and FRED. In appendix B.1 a detailed description of the exact source can be found. In order to remove long term trends from the data I employ a HP filter with a smoothing parameter equal to 1600.[11]

[10]Note that I do not describe the joint distribution of the impulse response function of the DSGE model and the VAR model. This is due to the fact that the sign restrictions are identical across a reasonable parameter space of the DSGE model.

[11]As suggested by Hodrick and Prescott (1997).

3.5.3 Specification of the identifying restriction and the prior distribution

I order the government expenditure shock similarly as in Mountford and Uhlig (2005) to come second after a business cycle shock; output, real wages, private investment, private consumption and hours worked are restricted to respond positively. By doing so the business cycle shock is assumed to explain most of the variance of the variables of the VMA model. The government expenditure shock is constructed to be orthogonal to the business cycle shock, i.e. most of the potential co-movements will be removed. The pre-announcement horizon is set to three quarters following the number suggested by Yang (2007). All variables except private consumption and real wages are restricted to exhibit the signs derived from the impulse response function of the DSGE model. The sign is allowed to switch, i.e. it can differ during the pre-announcement of the shock and after its realization.

The mean of the prior distributions of the parameters of the DSGE model is specified very closely to values used for calibration by Galí et al. (2007). One exception is the choice of η. In order to allow for a wider range of impulse responses, the mean is set to 7. The standard deviations of the prior distribution are chosen to ensure that the impulse response functions of the DSGE model cover a wide range of possibilities as depicted in Figure B.3. Not all parameters are estimated. The values used to calibrate those parameters are taken entirely from Galí et al. (2007), except for the ratio of real debt to GDP which is set to $\frac{\bar{b}}{\bar{y}} = 0.6$. Calibrated parameters include: $\beta = 0.99$, $\delta = 0.025$, $\alpha = 1/3$, $\frac{\bar{g}}{y} = 0.2$, $\phi_\pi = 1.5$ and $\mu_\pi = 1.2$.

3.5.4 Estimation Results

The smoothed Kalman filter maximum likelihood estimates of the VMA model variables are depicted in Figure B.4. The plot indicates that the time series is very well described by the VMA model.

The impulse response functions of the VMA model are shown in Figure B.5. The first subplot displays the announcement and the realization of the government consumption expenditure shock: the first three periods are restricted to zero followed by one positive period. Afterwards, government expenditures fall and even become negative. Given the fact that real debt responds positively over five quarters (even though the response is restricted for four quarters only), the decline in government expenditure gives rise to a policy rule in which government consumption expenditure reacts negatively on an increase in real debt - as suggested among others by Leeper and Yang (2008), Bohn (1991) and Corsetti, Meier, and Müller (2009).

The responses of output and hours worked in Figure B.5 display a significantly different qualitative behavior before and after the realization of the government consumption expenditure shock. Both react negatively during the announcement horizon and strongly positively after the realization of the shock. The response of private investment shows similar behavior, but the sign of the response changes after four (not after the announcement horizon of three quarters) from negative to positive. This implies that the restrictions derived from the DSGE model are negative over the complete restriction horizon.

Summarizing the estimated sign restrictions from the DSGE model, it can be stated that real debt is restricted to react positively over four periods, private investment to react negatively, while output and hours worked are restricted to respond negatively during the announcement period and positively after the realization.

These identifying restrictions imply the following results for private consumption and real wages. Private consumption displays significantly negative behavior throughout the announcement period and a very strong positive response afterwards. Real wages react significantly positively on impact, decrease throughout the announcement period and increase after the realization. During the fifth and sixth period the response of real wages is positive.

What is driving this result? One advantage of the methodology employed is that the estimated impulse response functions of the VMA model can be interpreted in economic terms by means of the DSGE model. Key for the interpretation are the estimation results of the structural parameters of the DSGE model. Before I interpret the results, I present and discuss the estimation results of the structural parameters and the corresponding impulse response function of the DSGE model.

The prior and posterior distribution is plotted in Figure B.7. Table B.1 gives an overview of the characteristics of the prior and posterior distribution. The plot and the table indicate that all parameters are well identified and that the estimates are within a very reasonable range. The estimated share of rule-of-thumb consumers (0.36 at the posterior mean) is in the range of the estimates obtained by Forni, Monteforte, and Sessa (2009) in studies using European data: $0.34 - 0.37$. The posterior mean of the calvo parameter ($\vartheta = 0.74$) is estimated slightly higher than in a recent study by Smets and Wouters (2007), who estimated it at about ($\vartheta = 0.66$). However, the relatively high estimate of the price stickiness parameter is consistent with a characteristic of the DSGE model pointed out by Furlanetto and Seneca (2009): the DSGE model employed obeys many real frictions, which are present in the DSGE model of Smets and Wouters (2007). Once those real rigidities were included, the resulting nominal frictions would have been estimated at lower values. The posterior mean estimate of the elasticity of labor with respect to wages ($\nu = 0.14$) is in line with, though slightly lower than the estimates obtained by Rotemberg and Woodford (1997) and Rotemberg and Woodford (1999). The value for the estimate of the elasticity of the investment to capital-capital ratio with respect to Tobins‘s Q ($\eta = 4.2$) is higher than the value calibrated by Galí et al. (2007) ($\eta = 1$), implying less investment adjustment costs. Given the response of government expenditure in Figure B.5, the persistence of government expenditure ($\rho_g = 0.74$) is lower than that calibrated by Galí et al. (2007) ($\rho_g = 0.9$). The elasticity of taxes with respect to government expenditure ($\phi_g = 0.08$) and the elasticity of taxes with respect to government debt ($\phi_b = 0.32$) are closely estimated to the values set by Galí et al. (2007) (($\phi_g = 0.1$) and $\phi_=0.33$).

A comparison of the impulse response functions of the DSGE model and the VMA model is plotted in Figure B.6. The impulse response functions of the VMA model are well matched by those of the DSGE model. In particular, the movement of output, hours worked, private consumption, and real wages (for some periods) are very well matched by the DSGE model. Impulse response functions not matched very well include private investment, government consumption expenditure and real debt after they become negative.

The impulse response of government consumption expenditure of the VMA model is very well matched by the DSGE model up to the point where government consumption expenditure becomes negative. Since the DSGE model employed in this analysis does not feature a policy rule with spending reversals, the behavior of government consumption expenditure after the shock cannot be matched by the impulse response function of the DSGE model. A policy rule featuring government spending reversals would be straightforward to implement. On the other hand, a modified policy rule would also affect other endogenous variables differently compared to the original DSGE model I want to consider. In order to be as transparent as possible on the choice of the DSGE model and its consequences, I choose to abstain from modifying the DSGE model in any dimension. A similar reasoning, i.e. a lack in the richness of the specification of the fiscal policy rules, applies to the response of real debt. Initially, the DSGE model has a lower response of real debt, which increases over time, while the impulse response function of the VMA model displays a strong initial positive response, which becomes negative after five periods.

The impulse response of private investment displays the largest difference between the VMA model and the DSGE model: the VMA model displays a strong negative response, which becomes positive after the restriction horizon of four periods. Even though, the adjustment costs of inflation are estimated more flexibly than in the original calibration by Galí et al. (2007), the response of investment is less volatile in the DSGE model than in the VMA model.

Most importantly, the impulse response functions of the DSGE model and the VMA model, which are at the center of the analysis, correspond very well: firstly the variables which display a qualitative difference between the announcement and the realization of the shock, i.e. output and hours worked; and secondly the variables under inspection, private consumption over the complete horizon and real wages except for the first period. Since the impulse responses correspond so well, the DSGE model can be used to recover the mechanism behind the impulse responses of private consumption and real wages. Key for the mechanism are the expectations formed by the firm sector. Figure B.8 shows the impulse response functions of the nominal interest rate and inflation in the DSGE model. Both responses rise after the pre-announcement of the shock. Firms anticipate higher inflation after the realization of the shock and increase prices after the announcement immediately. Monetary authorities increase nominal interest rates as a response to the rise in inflation. The increase in prices leads to an initial drop in GDP, hours worked, private consumption and investment. The result is stagflation during the announcement. The increase in government expenditure leads then to a rise in output, hours worked and private consumption for two reasons: a rise in income for the share of rule-of-thumb households and a drop in prices due to the negative response of output so far - prices overshoot.

3.5.5 Fiscal Multiplier and Variance decomposition

I compute two kinds of fiscal multipliers. The first multiplier is defined as the ratio of the response of output divided by the change in government expenditure and scaled by the average share of government consumption expenditure in GDP over the sample[12]. The result is shown in Figure B.9. Due to the negative response of output after the announcement, the fiscal multiplier is negative. After two periods it is strongly rising, becomes positive after four periods and significantly larger than 1 after five periods. Eventually it is decreasing again.

The second multiplier is defined as the cumulative change in output divided by the change in government consumption expenditure and scaled by the average share of

[12]This definition is also used in Mountford and Uhlig (2008).

government consumption expenditure in GDP over the sample. The result is depicted in Figure B.10. The negative response of output in the beginning causes the multiplier to not be positive until five periods after the announcement.

The share of the variance of the variables explained by the two shocks is shown in Figure B.11 and Figure B.12 for the business cycle and the government consumption expenditure shock respectively. As assumed, the business cycle shock explains most of the variance in all variables. For all variables except government expenditure this is up to 40% with a median around 10%. The government expenditure shock explains up to 10% of the variance during the announcement period for the variables except real wages. Here up to 40% is explained during the impact period, but much less, below 15%, in the following periods. After the realization of the shock, up to 20% of the variance in output and private consumption is explained.

3.5.6 Comparison with other studies

To the best of my knowledge there are three studies considering the effects of a pre-announced government expenditure shock using a SVAR model approach.

The results presented in this paper are most similar to those obtained by Mertens and Ravn (2009). In contrast to the other studies they also address the issue of the non-invertibility of the VMA representation. Even though they do not find qualitative differences in the response of private consumption and output after the announcement and the realization of a shock, they find a very strong announcement effect: both variables increase strongly when the shock takes place. During the announcement private consumption reacts negatively.

The results obtained are also in line with Tenhofen and Wolff (2007). They consider a one quarter announcement horizon and find a negative response of private consumption to a pre-announced government expenditure shock. After the shock, private consumption increases steadily, but, in contrast to the result in this paper, does not become positive.

Mountford and Uhlig (2008) consider different policy scenarios for a four quarter announcement horizon without explicitly modeling the pre-announcement. For an announced increase in government expenditure they find an immediate rise in private consumption and output as the effect of the announcement. Despite a more persistent response in output and private consumption they do not find any other announcement effects.

Putting the results into the context of the debate between Ramey (2008) and Blanchard and Perotti (2002), I find partial support for both views: while there are pre-announcement effects that cause private consumption to respond negatively during the first periods as pointed out by Ramey (2008), the realization of the shock leads to a strong positive response in private consumption as found by Blanchard and Perotti (2002).

3.6 Conclusion

This paper has investigated the effect of a government expenditure shock on private consumption and real wages by employing a structural VMA model. The identification key has been to model the pre-announcement of a government expenditure shock and its consequences on other economic variables explicitly.

The application of this idea is not straightforward for two reasons: first, when assuming that policy is pre-announced, the moving average representation of the data generated by this policy is potentially non-stable so that it cannot be approximated by a VAR model. I have therefore estimated a VMA model directly. Second, since the restrictions are not common knowledge I have employed a DSGE model, laid out initially by Galí et al. (2007), from which to derive the sign restrictions. The DSGE model is well suited to the problem because it addresses the typical arguments of the Keynesian as well as the classic view of the economy. On the one hand it features households which cannot smooth consumption, imperfect labor markets and a certain degree of price stickiness. How strong these features influence the result and the restrictions depend on its parametrization, for example the proportion of rule-of-thumb consumers and the degree of price stickiness. In the limit, i.e. with no rule-of-thumb consumers and firms allowed to reset prices each period, it boils down to a neoclassical model. Therefore, as Figure B.3, indicates this DSGE model allows for positive as well as negative responses in consumption and real wages. The parametrization of the DSGE model and the corresponding identifying assumptions for the VMA model are estimated by matching the corresponding impulse response functions of the VMA model. Thus the parameters of the VMA model and the DSGE model are estimated jointly.

The results for a three quarter pre-announced increase in government expenditures show strong qualitative differences during the announcement period and after the realization of the shock: output and hours worked respond negatively during the announcement period and positively afterwards, investment responds negatively one additional quarter before responding positively. Private consumption mimics this behavior and shows a stable, slightly negative response during the announcement period followed by a significant positive response after the realization of the shock. Real wages react significantly positively on impact, decrease (and even become negative) during the announcement horizon and react significantly positively for two quarters after the realization.

Chapter 4

Optimal Policy under Model Uncertainty: A Structural-Bayesian Estimation Approach

Uncertainty about the appropriate choice among nested models is a central concern for optimal policy when policy prescriptions from those models differ. The standard procedure is to specify a prior over the parameter space ignoring the special status of some sub-models, e.g. those resulting from zero restrictions. This is especially problematic if a model's generalization could be either true progress or the latest fad found to fit the data. We propose a procedure that ensures that the specified set of sub-models is not discarded too easily and thus receives no weight in determining optimal policy. We find that optimal policy based on our procedure leads to substantial welfare gains compared to the standard practice.

4.1 Introduction

Recently, the empirical evaluation of Dynamic Stochastic General Equilibrium (DSGE) models employing Bayesian methods has made substantial progress [1]. Policymakers nowadays correspondingly employ relatively large estimated DSGE models, including various features and frictions, in their policy analysis more and more. This practice is based on the implicit idea that by capturing many aspects of the economy in one single model, policy prescriptions derived from this model should guard against the risks of

[1](Smets and Wouters, 2003, 2007; An and Schorfheide, 2007; Lubik and Schorfheide, 2004).

an uncertain economic environment.[2] However, as it ignores the special status of sub-models that are defined by zero restrictions, the recent practice is prone to uncertainty about the appropriate choice of nested models. We show that this source of uncertainty is a central concern for optimal policy and propose a procedure that insures against it by assigning a non-zero weight to the set of sub-models.

To fix ideas, consider the following situation. After some process of theorizing and data analysis, a policymaker has arrived at a baseline model, Model A. One day, a researcher proposes to extend this model by adding a new feature or friction, replacing it with Model B, in which Model A is nested. At a first glance, this seems to be a win-win situation because the new model nests all the advantages of Model A and moreover may improve the understanding of the economy and lead the policymaker to make better policy decisions. However, the gain in explanatory power may be relatively small, i.e. the posterior odds may not indicate substantial evidence against Model A. Discarding Model A is further problematic because instead of true improvement, Model B may be just the latest fad found to fit the data. When Model B introduces a conflicting stabilization aim into the decision about policy, optimal policy prescriptions from the two models differ. In this situation, the policymaker risks welfare losses by ignoring Model A and putting all her eggs in one basket. In this chapter, we develop an approach that takes into account both Models A and B to determine optimal policy.

Starting with a baseline model, we subsequently estimate a set of competing and nested models. This bottom-up approach puts us into a position to separately evaluate the gain in explanatory power of each extension. Optimal policy is then computed by weighting each model with its posterior probability. Weighting over the set of nested models allows the policymaker to make reasonable extensions of the baseline model but also insure against the pitfalls of only employing one potentially misspecified model.

Using Euro-13 area data, we illustrate our approach to deal with model uncertainty in nested models by choosing as a baseline model one of the most popular models employed in monetary analysis nowadays: a standard cashless New Keynesian economy with staggered price-setting without indexation (Woodford, 2003a). As examples of uncertainty linked to the choice between nested models, we subsequently allow for more lags in endogenous variables (indexation and habit formation) and omitted variables (money). To represent the standard practice, we also consider one model that nests all these features. While the predominant principle of optimal policy in cashless models is price stability, a demand for money introduces a conflicting policy aim, namely

[2]Exemplary papers that fall in this category are Levin et al. (2005), Christiano, Trabandt, and Walentin (2007), and more recently Adolfson, Laseen, Linde, and Svensson (2008).

the stabilization of the nominal interest rate. In this environment, we find that our procedure leads to welfare gains of approximately 70 percent compared to the standard practice.

The remainder of the chapter is organized as follows. In the next section we introduce our approach to analyze the optimal conduct of policy under model uncertainty. In Section 4.3 we describe the baseline model and its extensions. In Section 4.4 we present our estimation results and its consequences for optimal monetary policy. The last section concludes.

4.2 Analyzing optimal policy under model uncertainty

In this section, after a short description of the general setup we present two approaches to cope with model uncertainty and describe how we assess the policy performance under model uncertainty. The first approach is set to represent the standard practice: without paying special attention to the set of sub-models, the policymaker determines optimal policy by maximizing households' utility within one single model that nests all features and frictions. The second approach takes uncertainty about the appropriate choice of nested models into account and weights over the set of nested models to derive optimal policy prescriptions.

4.2.1 General setup

Consider a system of linear equations that represent log-linear approximations to the non-linear equilibrium conditions under rational expectations around a deterministic steady state of a particular Model i. Let x_t be the vector of state variables, z_t the vector of structural shocks and y_t the vector of observable variables. Furthermore, let Θ_i denote the random vector of deep parameters and θ_i a particular realization from the joint posterior distribution in Model i. Policy influences the equilibrium outcome through simple feedback rules. The link between the set of policy instruments as a subset of x is characterized by the vector of constant policy coefficients ϕ, i.e. by definition we consider steady state invariant policies.[3] The state space form of the

[3] A steady state-invariant policy is a policy which affects the dynamic evolution of the endogenous variables around a steady state, but not the steady state itself.

solution of model i is given by[4]:

$$\widehat{x}_t = T(\theta_i, \phi)\widehat{x}_{t-1} + R(\theta_i, \phi)z_t \tag{4.1}$$

$$\widehat{y}_t = G\widehat{x}_t, \tag{4.2}$$

where $T(\theta_i, \phi)$ and $R(\theta_i, \phi)$ are matrices one obtains after solving a DSGE model with standard solution techniques. The matrix G is a picking matrix that equates observable and state variables.

We assess the performance of a particular policy ϕ by its effects on households' unconditional expected utility, i.e. before any uncertainty has been resolved. In Model i and for a particular realization θ_i, this unconditional expectation up to second order is represented by:

$$E\sum_{t=t_0}^{\infty} \beta^t U(x_t, \theta_i) \approx \frac{U(\bar{x}, \theta_i)}{1-\beta} - E\sum_{t=t_0}^{\infty} \beta^t A(\theta_i)\widehat{x}_t\widehat{x}_t' = \frac{U(\bar{x}, \theta_i)}{1-\beta} - \frac{L(\theta_i, \widehat{x})}{1-\beta}. \tag{4.3}$$

This approximation decomposes households' utility in two parts. The first part is utility in the steady state, and the second part comprises welfare-reducing fluctuations around the long-run equilibrium. We assume that the policymaker can credibly commit to a policy rule ϕ: if a policymaker decides to follow a certain policy rule ϕ once and forever, agents believe indeed that the policymaker will. Given a particular value θ_i, the optimal steady state invariant policy $\phi_i^\star(\theta_i)$ maximizes (4.3) by minimizing short-run fluctuations captured in $L(\theta_i, \widehat{x})$. Since the specification of households' preferences is independent of policy choices, the policymaker can only indirectly influence households' loss by shaping the dynamics of the endogenous variables $\widehat{x}$ as defined by (4.1).

4.2.2 Two approaches to model uncertainty

We now turn to the optimal conduct of policy if the policymaker faces uncertainty about the economic environment. We consider two approaches to cope with this uncertainty.

Specifying a marginal prior distribution with a positive unique mode for each parameter, the first approach or the standard practice is to develop and estimate one single model that nests all features and frictions and employ the model in determining optimal policy. This is based on the idea that by capturing many aspects of the economy in one single model, policy prescriptions derived from this model should guard

[4] $\widehat{x}_t$ denotes the percentage deviation of the generic variable x_t from a deterministic steady state x chosen as approximation point.

against the risks of an uncertain economic environment. The only source of uncertainty for the policymaker is uncertainty about the structural parameters of the model. We refer to this approach as the *complete-model* approach.

The second approach starts with a stylized baseline model and treats each extension by an additional feature or friction as a distinct and competing model. By averaging across models, this approach allows to take not only parameter uncertainty but also uncertainty about model specification into account. In the following we refer to this approach as the *model-averaging* approach.

When pursuing the first approach to deal with model uncertainty, the relevant uncertainty that a policymaker faces when she makes her decision about ϕ is given by the joint posterior distribution in the model that nests all features and frictions. We denote this 'complete' model by Model $\textsc{c}$ and its corresponding posterior distribution of its structural parameters by $f(\theta_c) \equiv f(\theta|Y, \mathcal{M}_{\textsc{c}})$, where Y is the set of time series used in the estimation. The optimal policy ($\phi^\star_{\textsc{c}}$) is defined by:

$$\begin{aligned} \phi^\star_{\textsc{c}} &= \arg\min_{\phi} E_{\Theta_{\textsc{c}}} L(\Theta_{\textsc{c}}, \widehat{x}) \\ s.t.\ \widehat{x}_t &= T(\theta_{\textsc{c}}, \phi)\widehat{x}_{t-1} + R(\theta_{\textsc{c}}, \phi) z_t, \quad \forall \theta_{\textsc{c}}, \end{aligned} \tag{4.4}$$

where $E_{\Theta_{\textsc{c}}} L(\Theta_{\textsc{c}}, \widehat{x})$ is the expected loss when the structural parameters are a random vector. Due to parameter uncertainty the policymaker has to average the loss over all possible realizations of $\Theta_{\textsc{c}}$ to find the optimal vector of constant policy coefficients in Model $\textsc{c}$, $\phi^\star_{\textsc{c}}$.

The second approach explicitly addresses specification uncertainty and averages over different models. We separately estimate a discrete set of nested models $\mathcal{M} = \{\mathcal{M}_1, ..., \mathcal{M}_{\textsc{c}}\}$, where $\mathcal{M}_1$ denotes the baseline model, $\mathcal{M}_{\textsc{c}}$ the complete model and $(\textsc{c} - 2)$ possible one-feature extensions of the baseline model. Employing the same data and prior specification of shocks and common parameters, we calculate marginal data densities $p(Y|\mathcal{M}_i) = E(f(Y|\Theta_i))$, where $f(Y|\Theta_i)$ denotes the data likelihood and the expectation is taken with respect to the prior distribution of the structural parameters. Since all models are nested in Model $\textsc{c}$, the marginal data density for Model i satisfies:

$$p(Y|\mathcal{M}_i) \equiv p(Y|\theta_{\textsc{c}\notin i} = \mathbf{0}, \mathcal{M}_{\textsc{c}}), \tag{4.5}$$

where $\theta_{\textsc{c}\notin i}$ denotes the vector of structural parameters for Model $\textsc{c}$ that are not contained in the set of structural parameters of Model $i = 1, 2, ...\textsc{c}$. We employ the harmonic mean estimator to compute the data likelihood in a certain model as proposed by Geweke (1999b) and more recently applied among others by An and Schorfheide (2007). To compare the explanatory power of each model relative to the other models, we compute posterior probabilities which are defined as

$$P(M_i|Y) = \frac{P(M_i)p(Y|\mathcal{M}_i)}{\sum_{j=1}^{\mathrm{C}} P(M_j)p(Y|\mathcal{M}_j)}, \tag{4.6}$$

where $P(M_i)$ denotes the prior probability for each model.[5] To ensure that sub-models are not discarded too easily and to facilitate competition between the nested models, we assign positive and equal prior weights to each model. The posterior probability of each model is then solely determined by its relative success to explain a given set of time series, i.e. it takes a value close to zero when the predictive density of a model relative to the others is neglectable.

In nested models, the second approach can also be thought of as defining a bimodal prior distribution for the parameter that represents the additional feature or friction. One part of the distribution is centered around the assumed positive modulus of the parameter, and the other modulus is centered around zero. The idea of paying special attention to this zero restriction and giving this possibility relatively more weight reflects the natural scepticism every researcher and policymaker has when extending a reasonable model.

Our approach however is more general than specifying bimodal prior distributions because it can also be applied when models are not nested. In particular, it avoids a discontinuity problem in the parameter space that arises when models are not nested. To see this, suppose that the baseline model $\mathcal{M}_1$ is replaced by a very similar model $\mathcal{M}_1^*$ that is not nested in the complete Model C. In other words, there is at least one parameter that is not included in the prior specification of the complete model. In this case, it seems to be reasonable to weight over all models, also including $\mathcal{M}_1^*$. The complete-model approach – even if it includes a bimodal prior specification – gives zero weight to the parameter included in $\mathcal{M}_1^*$ but not in Model C. The model-averaging approach weights over models independent whether they are nested or not, and thereby avoids this discontinuity. In addition, the formulation of a bimodal prior distribution in standard Bayesian model estimation is not straightforward and estimating a model extension to zero might cause serious troubles when approximating the posterior mean.

[5] An alternative approach to compute posterior model probabilities in nested models involves calculating Savage-Dickey density ratios as proposed by Verdinelli and Wasserman (1995).

The optimal policy for the model-averaging approach ($\phi_a^\star$) is defined by

$$\phi_a^\star = \arg\min_{\phi} E_{\mathcal{M},\Theta} L(\Theta_i, \widehat{x}) \tag{4.7}$$
$$s.t.\ \widehat{x}_t = T(\theta_i, \phi)\widehat{x}_{t-1} + R(\theta_i, \phi) z_t, \quad \forall\theta_i, \quad i = 1, ..., n.$$

The complete-model approach is a limiting case of model-averaging approach; they are equivalent if the complete model exhibits a posterior probability of unity.

4.2.3 Assessing policy performance within and across models

We compare the performance of the two approaches by computing the average costs of welfare relevant short-run fluctuations over all draws and models. This allows us to assess the pitfalls of employing only one model that nests all features and frictions in the policy analysis, i.e. focussing on parameter uncertainty in the complete model and thereby ignoring the issue of specification uncertainty about nested models. Throughout the chapter we express the resulting business cycle costs ($\mathcal{BC}$) as the percentage loss in certainty (steady state) equivalent consumption. First we compute the loss of a certain policy $\tilde{\phi}$ given a particular parameter vector $\tilde{\theta}$ in model i to derive overall utility:

$$U\left(c(\tilde{\theta}_i), x_{\setminus c}(\tilde{\theta}_i), \tilde{\theta}_i\right) - L(\tilde{\theta}_i, \tilde{\phi}),$$

where the first term is steady state utility and $x_{\setminus c}$ denotes the variables vector excluding consumption. Since we want to express utility as reduction in certainty consumption equivalents we set this expression to be equal to:

$$U\left(c(\tilde{\theta}_i)(1 - \mathcal{BC}), x_{\setminus c}(\tilde{\theta}_i), \tilde{\theta}_i\right)$$

and solve for $\mathcal{BC}$ in percentage terms. Under parameter uncertainty this results in a distribution for $\mathcal{BC}(\tilde{\theta}_i, \tilde{\phi})$ over Θ_i. Taking the expectation of this expression yields a measure of the average losses in certainty consumption equivalents under a particular policy $\tilde{\phi}$.

As can be seen from (4.3), theoretical unconditional second moments derived from the DSGE model are relevant for households' utility losses due to short fluctuations – and thus for the computation of business-cycle costs under different policies. As Del Negro and Schorfheide (2008) point out, whether the theoretical unconditional moments relevant for policy assessment and the ones observed in the data coincide depends in particular on the specification of the prior distribution of standard deviations and autoregressive coefficients for the driving exogenous disturbances. We choose the prior distribution for the standard deviations of the *i.i.d.* terms in z_t and the autoregressive coefficients of the shocks contained in $T(\bullet)$ such that the relevant theoretical unconditional second moments at the posterior mean in each model are in line the ones computed directly from the stationary times series. This in turn yields welfare costs of short run fluctuations consistent with the limit put forward by Lucas (2003).

4.3 Optimal monetary policy: the economic environment

To demonstrate our main result, we create a set of monetary models including one model that nests all features and frictions. Starting with a plain-vanilla cashless new Keynesian economy as our baseline model (Woodford, 2003a), we subsequently introduce two additional features (indexation and habit formation) and a transaction friction (money in the utility function). While optimal policy in the baseline model and in the models that feature indexation and habit formation seeks to stabilize fluctuations in inflation and in the output gap, a transaction friction adds the stabilization of the nominal interest rate as an additional and conflicting policy aim. In this section we describe the models, derive the equations characterizing the equilibrium and the relevant policy objectives as the unconditional expectation of households' utility for each model.

4.3.1 The baseline economy: Model 1

The baseline economy consists of a continuum of infinitely-lived households indexed with $j \in [0,1]$ that have identical initial asset endowments and identical preferences. Household j acts as a monopolistic supplier of labor services l_j. Lower (upper) case letters denote real (nominal) variables. At the beginning of period t, households' financial wealth comprises a portfolio of state contingent claims on other households yielding a (random) payment Z_{jt}, and one-period nominally non-state contingent government bonds B_{jt-1} carried over from the previous period. Assume that financial markets are complete, and let $q_{t,t+1}$ denote the period t price of one unit of currency in a particular state of period $t+1$ normalized by the probability of occurrence of that state, conditional on the information available in period t. Then, the price of a random payoff Z_{t+1} in period $t+1$ is given by $E_t[q_{t,t+1}Z_{jt+1}]$. The budget constraint of the representative household reads

$$B_{jt} + E_t[q_{t,t+1}Z_{jt+1}] + P_t c_{jt} \leq R_{t-1}B_{jt-1} + Z_{jt} + P_t w_{jt} l_{jt} + \int_0^1 D_{jit} di - P_t T_t, \quad (4.8)$$

where c_t denotes a Dixit-Stiglitz aggregate of consumption with elasticity of substitution ζ, P_t the aggregate price level, w_{jt} the real wage rate for labor services l_{jt} of type j, T_t a lump-sum tax, R_t the gross nominal interest rate on government bonds, and D_{it} dividends from monopolistically competitive firms. The objective of the representative household is

$$E_{t_0} \sum_{t=t_0}^{\infty} \beta^t \{u(c_{jt}) - v(l_{jt})\}, \quad \beta \in (0,1), \quad (4.9)$$

where β denotes the subjective discount factor. The instantaneous utility function is assumed to be non-decreasing in consumption, decreasing in labor time, strictly concave, twice differentiable, and to fulfill the Inada conditions. Households are wage-setters supplying differentiated types of labor l_j, which are transformed into aggregate labor l_t with $l_t^{(\epsilon_t-1)/\epsilon_t} = \int_0^1 l_{jt}^{(\epsilon_t-1)/\epsilon_t} dj$. We assume that the elasticity of substitution between different types of labor, $\epsilon_t > 1$, varies exogenously over time. Cost minimization implies that the demand for differentiated labor services l_{jt}, is given by $l_{jt} = (w_{jt}/w_t)^{-\epsilon_t} l_t$, where the aggregate real wage rate w_t is given by $w_t^{1-\epsilon_t} = \int_0^1 w_{jt}^{1-\epsilon_t} dj$. The transversality condition is given by

$$\lim_{i\to\infty} E_t \beta^i \lambda_{jt+i}(B_{jt+i} + Z_{jt+1+i})/P_{jt+i} = 0 \quad (4.10)$$

. The final consumption good Y_t is an aggregate of differentiated goods produced by monopolistically competitive firms indexed with $i \in [0,1]$ and defined as $y_t^{\frac{\zeta-1}{\zeta}} = \int_0^1 y_{it}^{\frac{\zeta-1}{\zeta}} di$, with $\zeta > 1$. Let P_{it} and P_t denote the price of good i set by firm i and the price index for the final good. The demand for each differentiated good is $y_{it}^d = (P_{it}/P_t)^{-\zeta} y_t$, with $P_t^{1-\zeta} = \int_0^1 P_{it}^{1-\zeta} di$. A firm i produces good y_i using a technology that is linear in the labor bundle $l_{it} = [\int_0^1 l_{jit}^{(\epsilon_t-1)/\epsilon_t} dj]^{\epsilon_t/(\epsilon_t-1)}$: $y_{it} = a_t l_{it}$, where $l_t = \int_0^1 l_{it} di$ and a_t is a productivity shock with mean 1. Labor demand satisfies: $mc_{it} = w_t/a_t$, where $mc_{it} = mc_t$ denotes real marginal costs independent of the quantity that is produced by the firm. We allow for a nominal rigidity in form of a staggered price setting as developed by Calvo (1983). Each period firms may reset their prices with probability $1-\alpha$ independently of the time elapsed since the last price setting. A fraction $\alpha \in [0,1)$ of firms are assumed to keep their previous period's prices, $P_{it} =$

P_{it-1}. In each period a measure $1-\alpha$ of randomly selected firms set new prices $\widetilde{P}_{it}$ as the solution to

$$\max_{\widetilde{P}_{it}} E_t \sum_{T=t}^{\infty} \alpha^{T-t} q_{t,T}(\widetilde{P}_{it} y_{iT}(1-\tau) - P_T mc_T y_{iT}), \quad s.t.\ y_{iT} = (\widetilde{P}_{it})^{-\zeta} P_T^{\zeta} y_T, \tag{4.11}$$

where τ denotes an exogenous sales tax. We assume that firms have access to contingent claims.

The aggregate resource constraint is given by

$$y_t = a_t l_t / \Delta_t, \tag{4.12}$$

where $\Delta_t = \int_0^1 (P_{it}/P_t)^{-\zeta} di \geq 1$ and thus $\Delta_t = (1-\alpha)(\widetilde{P}_t/P_t)^{-\zeta} + \alpha \pi_t^{\zeta} \Delta_{t-1}$. The dispersion measure Δ_t captures the welfare decreasing effects of staggered price setting. Goods' market clearing requires

$$c_t + g_t = y_t. \tag{4.13}$$

The central bank as the monetary authority is assumed to control the short-term interest rate R_t with a simple feedback rule contingent on past interest rates, inflation and output:

$$R_t = f(R_{t-1}, \pi_t, y_t). \tag{4.14}$$

The consolidated government budget constraint reads: $R_{t-1}B_{t-1} + P_t G_t = B_t + P_t T_t + \int_0^1 P_{it} y_{it} \tau di$. The exogenous government expenditures g_t evolve around a mean $\bar{g}$, which is restricted to be a constant fraction of output, $\bar{g} = \bar{y}(1-sc)$. We assume that tax policy guarantees government solvency, i.e., ensures $\lim_{i\to\infty} (B_{t+i}) \prod_{v=1}^{i} R_{t+v}^{-1} = 0$.

We collect the exogenous disturbances in the vector $\xi_t = [a_t, g_t, \mu_t]$, where $\mu_t = \frac{\epsilon_t}{\epsilon_t - 1}$ is a wage mark-up shock. It is assumed that the percentage deviations of the first two elements of the vector from their means evolve according to autonomous AR(1)-processes with autocorrelation coefficients $\rho_a, \rho_g \in [0,1)$. The process for $\log(\mu_t/\bar{\mu})$ and all innovations, $z_t = [\varepsilon_t^a, \varepsilon^{\tilde{g}_t}, \varepsilon_t^{\mu}]$, are assumed to be i.i.d..

The recursive equilibrium is defined as follows:

Definition 3 *Given initial values $P_{t_0-1} > 0$ and $\Delta_{t_0-1} \geq 1$, a monetary policy and a Ricardian fiscal policy $T_t\ \forall\, t \geq t_0$, and a sales tax τ, a rational expectations equilibrium (REE) for $R_t \geq 1$, is a set of sequences $\{y_t,\ c_t,\ l_t,\ mc_t,\ w_t,\ \Delta_t,\ P_t,\ \widetilde{P}_{it}, R_t\}_{t=t_0}^{\infty}$ for $\{\xi_t\}_{t=t_0}^{\infty}$*

(i) that solve the firms' problem (4.11) with $\widetilde{P}_{it} = \widetilde{P}_t$,

(ii) that maximize households' utility (4.9) s.t. their budget constraints (4.8),

(iii) that clear the goods market (4.13),

(iv) and that satisfy the aggregate resource constraint (4.12) and the transversality condition (4.10).

In the next step, we seek to estimate the model by employing Bayesian methods. To do so, we log-linearize the structural equations around the deterministic steady state under zero inflation. Thus, the dynamics in the baseline economy are described by the following two structural equations:

$$\sigma(E_t\widehat{y}_{t+1} - E_t\widehat{y}^n_{t+1}) = \sigma(\widehat{y}_t - \widehat{y}^n_t) + \widehat{R}_t - E_t\widehat{\pi}_{t+1} - \widehat{R}^n_t \tag{4.15}$$

$$\widehat{\pi}_t = \beta E_t\widehat{\pi}_{t+1} + \kappa(\widehat{y}_t - \widehat{y}^n_t), \tag{4.16}$$

where $\sigma = -u_{cc}c/(u_c sc)$, $\omega = v_{ll}l/v_l$ and $\kappa = (1-\alpha)(1-\alpha\beta)(\omega+\sigma)/\alpha$. Furthermore, $\widehat{k}_t$ denotes the percentage deviation of a generic variable k_t from its steady-state value k. The natural rates of output and interest, i.e the values for output and real interest under flexible prices, are given by the following expressions

$$\widehat{y}^n_t = \frac{(1+\omega)\widehat{a}_t + \sigma\tilde{g}_t - \widehat{\mu}_t}{\omega+\sigma}, \quad \widehat{R}^n_t = \sigma[(\tilde{g}_t - \widehat{y}^n_t) - E_t(\tilde{g}_{t+1} - \widehat{y}^n_{t+1})],$$

where $\tilde{g}_t = (g_t - g)/y$. The model is closed by a simple interest rate feedback rule as an approximation to (4.14):

$$\widehat{R}_t = \rho_R\widehat{R}_{t-1} + \phi_\pi\widehat{\pi}_t + \phi_y\widehat{y}_t. \tag{4.17}$$

The general system (4.1) in the baseline model then is the fundamental locally stable and unique solution that satisfies (4.15)-(4.17) for a certain vector of constant policy coefficients $\phi = (\rho_R, \phi_\pi, \phi_y)$.

Our welfare measure is the unconditional expectation of representative households' utility. Building on Woodford (2003a), after averaging over all households, a second-order approximation to (4.9) results in the following quadratic loss function (for a given realization θ_1)[6]:

$$L(\theta_1, \widehat{x}) = \frac{u_c y\zeta(\omega+\sigma)}{2\kappa}\{var(\widehat{\pi_t}) + \lambda_d var(\widehat{y}_t - \widehat{y}^e_t)\}, \tag{4.18}$$

[6]Throughout we assume that the steady state is rendered efficient by an appropriate setting of the sales tax rate.

where $\lambda_d = \kappa/\zeta$ and the efficient rate of output is given by

$$\widehat{y}_t^e = \widehat{y}_t^n + \widehat{\mu}_t/(\omega + \sigma).$$

In the next subsection we consider habit formation and indexation to past inflation as examples of missing lags in consumption and inflation.

4.3.2 Habit formation (Model 2) and indexation (Model 3)

One example of a missing lag in an endogenous variable is to allow for an internal habit (e.g. Boivin and Giannoni (2006); Woodford (2003a)) in households' total consumption. The constituting equations for (4.1) are the policy rule (4.17) and the modified versions of the Euler equation and the New Keynesian Philips curve:

$$\begin{aligned}\varphi[d_t - \eta d_{t-1}] - \varphi\beta\eta E_t[d_{t+1} - \eta d_t] = E_t\widehat{\pi}_{t+1} + \widehat{R}_t^n - \widehat{R}_t ...\\ + E_t\varphi[d_{t+1} - \eta d_t] - \varphi\beta\eta E_t[d_{t+2} - \eta d_{t+1}] \quad (4.19)\end{aligned}$$

$$\widehat{\pi}_t = \kappa_h[(d_t - \delta^* d_{t-1}) - \beta\delta^* E_t(d_{t+1} - \delta^* d_t)] + \beta E_t\widehat{\pi}_{t+1}, \quad (4.20)$$

where $d_t = \widehat{y}_t - \widehat{y}_t^n$, $\kappa_h = \eta\varphi\kappa[\delta^*(\omega + \sigma)]^{-1}$, $\varphi = \sigma/(1 - \eta\beta)$, and the natural rate of output follows[7]

$$\begin{aligned}[\omega + \varphi(1 + \beta\eta^2)]\widehat{y}_t^n - \varphi\eta\widehat{y}_{t-1}^n - \varphi\eta\beta E_t\widehat{y}_{t+1}^n = \varphi(1 + \beta\eta^2)\tilde{g}_t - \varphi\eta\tilde{g}_{t-1} - \varphi\eta\beta E_t g_{t+1} ...\\ + (1 + \omega)\widehat{a}_t - \widehat{\mu}_t.\end{aligned}$$

Approximating households' utility to second order results in the following loss function:

$$L(\theta_2, \widehat{x}) = \frac{(1 - \beta\eta)\eta\varphi u_c^h y^h \zeta}{2\kappa_h\delta^*}\{var(\widehat{\pi}_t) + \lambda_{d,h} var(\widehat{y}_t - \widehat{y}_t^e - \delta^*(\widehat{y}_{t-1} - \widehat{y}_{t-1}^e))\}, \quad (4.21)$$

where $\lambda_{d,h} = \kappa_h/\zeta$ and the efficient rate of output is characterized by

$$\begin{aligned}[\omega + \varphi(1 + \beta\eta^2)]\widehat{y}_t^e - \varphi\eta\widehat{y}_{t-1}^e - \varphi\eta\beta E_t\widehat{y}_{t+1}^e = \varphi(1 + \beta\eta^2)\tilde{g}_t - \varphi\eta\tilde{g}_{t-1} - \varphi\eta\beta E_t g_{t+1} ...\\ + (1 + \omega)\widehat{a}_t.\end{aligned}$$

[7]The parameter δ^*, $0 \leq \delta^* \leq \eta$, is the smaller root of the quadratic equation $\eta\varphi(1 + \beta\delta^2) = [\omega + \varphi(1 + \beta\eta^2)]\delta$. This root is assigned to past values of the natural and efficient rate of output in their stationary solutions.

Like habit formation, the indexation of prices to past inflation induces the economy to evolve in a history-dependent way. We assume that the fraction of prices that are not reconsidered, α, adjusts according to $\log P_{it} = \log P_{it-1} + \gamma \log \pi_{t-1}$ with $0 \leq \gamma \leq 1$ as the degree of indexation. This implies that price dispersion evolves according to $\Delta_t = (1-\alpha)(\frac{\widetilde{P}_t}{P_t})^{-\zeta} + \alpha \pi_{t-1}^{-\zeta\gamma} \Delta_{t-1} \pi_t^{\zeta}$. Correspondingly, the economy with indexation is characterized by a modified aggregate supply curve

$$\widehat{\pi}_t - \gamma\widehat{\pi}_{t-1} = \beta E_t(\widehat{\pi}_{t+1} - \gamma\widehat{\pi}_t) + \kappa(\widehat{y}_t - \widehat{y}_t^n), \tag{4.22}$$

(4.15) and (4.17). The corresponding loss function of the central bank reads Woodford (2003a):

$$L(\theta_3, \widehat{x}) = \frac{u_c y \zeta(\omega + \sigma)}{2\kappa}\{var(\widehat{\pi}_t - \gamma\widehat{\pi}_{t-1}) + \lambda_d var(\widehat{y}_t - \widehat{y}_t^e)\}, \tag{4.23}$$

where λ_d and the efficient rate of output are defined as in the baseline economy.

4.3.3 Money in the utility function (Model 4)

We introduce a transaction friction by letting real money balances enter households' utility in a separable way. More precisely, households' utility of holding real money balances is augmented by the amount $z(m_t)$ and a demand equation for real money balances enters the set of equilibrium conditions. In log-linearized form this additional equilibrium condition is given by:

$$\widehat{m}_t = -\frac{1}{\sigma_m(R-1)}\widehat{R}_t - \frac{1}{\sigma_m}\widehat{\lambda}_t, \tag{4.24}$$

where $\sigma_m = -z_{mm}m/z_m$ and $\hat{\lambda}_t$ denotes the Lagrangian multiplier on the budget constraint of the household. The stabilization loss in Model 4 is given by:

$$L(\theta_4, \widehat{x}) = \frac{u_c y \zeta(\omega + \sigma)}{2\kappa}\{var(\widehat{\pi}_t) + \lambda_d var(\widehat{y}_t - \widehat{y}_t^e) + \lambda_{1R} var(\widehat{R}_t)\}, \tag{4.25}$$

where $\lambda_d = \kappa/\zeta$, $\lambda_{1R} = \lambda_d \beta[v(\omega + \sigma)(1-\beta)\sigma_m]^{-1}$ and $v = y/m$. The general form (4.1) has to satisfy the (4.15)-(4.17) and (4.24).

4.3.4 The complete model

The complete model (Model C) builds on the baseline model and comprises habit formation, indexation and money in the utility function. The equilibrium conditions in this case are: (4.19), (4.17), (4.24) and

$$\widehat{\pi}_t - \gamma\widehat{\pi}_{t-1} = \beta E_t(\widehat{\pi}_{t+1} - \gamma\widehat{\pi}_t) + \kappa_h[(d_t - \delta^* d_{t-1}) - \beta\delta^* E_t(d_{t+1} - \delta^* d_t)]. \tag{4.26}$$

In the following proposition we state the loss function for Model C.

Proposition 1 *If the fluctuations in y_t around y, R_t around R, ξ_t around ξ, π_t around π are small enough, $(R-1)/R$ is small enough, and if the steady state distortions ϕ vanish due to the existence of an appropriate subsidy τ, the utility of the average household can be approximated by:*

$$U_{t_0} = -E_{t_0}\sum_{t=t_0}^{\infty}\beta^{t-t_0}L(\theta_{\text{C}}, \widehat{x}) + t.i.s.p. + \mathcal{O}(\|\widehat{\xi}_t, (R-1)/R\|^3), \tag{4.27}$$

where t.i.s.p. indicate terms independent of stabilization policy,

$$L(\theta_{\text{C}}, \widehat{x}) = \frac{(1-\beta\eta)\eta\varphi u_c^h y^h \zeta}{2\kappa_h\delta^*} \tag{4.28}$$
$$\{var(\widehat{\pi}_t - \gamma\widehat{\pi}_{t-1}) + \lambda_{d,h}var(\widehat{y}_t - \widehat{y}_t^e - \delta^*(\widehat{y}_{t-1} - \widehat{y}_{t-1}^e)) + \lambda_{2R}var(\widehat{R}_t)\},$$

$$\lambda_x = \lambda_{d,h} = \kappa_h/\zeta, \quad \lambda_{2R} = \frac{\lambda_{d,h}\beta\delta^*}{v\sigma_m(1-\beta)\eta\varphi},$$

and $v = y/m > 0$.

Proof: see appendix C.1.

4.4 Results

In this section we first present and interpret the estimation results. These results will be key for the assessment of the relevant model uncertainty faced by the policymaker. In the second part we compute optimal simple rules along with the procedures laid out in section 4.2. As a standard, we determine optimal monetary policy at the posterior mean, i.e. optimal policy in the absence of any model uncertainty. Then we analyze optimal policy when there is uncertainty about the appropriate choice of nested models.

4.4.1 Data and estimation results

We treat the variables real wage, output and consumer price inflation as observable. The data consists of HP filtered quarterly values of these variables for the EU 13 countries from 1970-2006.[8]

We calibrate the discount factor to $\beta = 0.99$, the steady-state fraction of private consumption relative to GDP $c/y = 0.8$ and the elasticity of substitution between differentiated goods to $\zeta = 6$ (see Woodford, 2003a). The specification of the prior distributions of the estimated deep parameters closely follows Negro and Schorfheide (2009), Smets and Wouters (2003) and Smets and Wouters (2007).[9] While we assume the disturbances $\tilde{g}_t$ and $\widehat{a}_t$ to follow stationary $AR(1)$ processes, $\widehat{\mu}_t$ is supposed to be *i.i.d.*. Since we are interested in evaluating the explanatory power of each extension of the baseline model separately, common parameters in the set of models need to exhibit the same sufficient prior statistics. In particular, the marginal prior distributions for the set of coefficients that describe the shock processes, $\psi_{\tilde{g}}, \psi_a$ and $\sigma_{\tilde{g}}, \sigma_a, \sigma_\mu$, do not change across models, and they are specified according to the procedure explained in Section 4.2.3.

We approximate the joint posterior distribution of structural parameters by drawing $100,000$ times employing a standard MCMC-algorithm as described in An and Schorfheide (2007) and discard the first $80,000$ draws. The estimation results are displayed in Table C.2 and the posterior and prior distributions are plotted in Figure 1-5 in Appendix C.2. The estimates of the posterior mean of the degree of relative risk aversion with respect to consumption (σ_c), the degree of indexation (γ), and the degree of price stickiness (α) correspond almost one-for-one to the findings by Smets and Wouters (2003). In line with Woodford (2003a) we find the labor supply decision with respect to changes in the real wage ($1/\omega$) to be elastic, i.e. values for ω vary between 0.3 and 0.4. Our estimate of the internal habit parameter (η) is comparable to Negro and Schorfheide (2009). Real money balances contribute only separately to households' utility in Model 4 and Model C and do not influence the equilibrium dynamics of output, inflation and the real wage. The parameter of relative risk aversion with respect to real money balances (σ_m) cannot be identified and thus the prior distribution and the posterior distribution are alike (see Figures 4 and 5).

In order to assess the explanatory power of each model, we compute marginal likelihoods and the corresponding posterior probabilities. The results are presented in

[8]The dataset we use was kindly provided by the Euro Area Business Cycle Network (EABCN). For a description of how this data is constructed see Fagan, Henry, and Mestre (2001).

[9]See Appendix C.2 Table C.1 for a detailed description.

Table 4.1. Here the key result is, that adding frictions and features to the baseline model does not necessarily increase the posterior probability. First, enriching the baseline model with a demand for cash does not increase the marginal likelihood for Model 4: real money balances do not help to predict the observable variables. Second, although the posterior distribution of the habit parameter (η) in Model 2 indicates a positive posterior mean of this parameter, a habit in consumption does not improve the fit to the data. This points to a well-known problem in Bayesian model estimation: The informative prior on the habit parameter introduces curvature into the posterior density surface (as pointed out by Poirier (1998) and An and Schorfheide (2007)). Third, history dependence in inflation improves the fit of the model. With approximately 81% Model 3 exhibits the highest posterior probability. Thus, the complete model incorporates features that helps to predict the data (indexation) and others that do not (habit and money). It therefore exhibits a marginal likelihood higher than Model 1 but lower than Model 3.

Table 4.1: Posterior probabilities and marginal data densities

	M_1	M_2	M_3	M_4	M_C
$p(Y\|\mathcal{M}_i)$	1683.98	1682.69	1696.83	1683.57	1695.39
$P(M_i\|Y)$	0.00	0.00	0.81	0.00	0.19

The welfare-assessment of optimal and sub-optimal policies in and across models depends on the magnitude of the resulting stabilization losses, i.e. the welfare relevant unconditional variances or standard deviations. In our context, these are the unconditional fluctuations in inflation and consumption (expressed in terms of a welfare-relevant output gap) for the models without a transaction friction (see e.g. (4.18)), and additionally fluctuations in interest rates, when money enters the utility function (see e.g (4.28)). As can be verified in Table 4.2, our estimated theoretical moments at the posterior mean are consistent with the corresponding ones directly estimated from the stationary times series.

Table 4.2: Welfare-relevant standard deviations: models vs. data

	M_1	M_2	M_3	M_4	M_C	*Data*
$std(c, \bar{\theta}_i)$	0.0070	0.0090	0.0068	0.0070	0.0078	0.0073
$std(\pi, \bar{\theta}_i)$	0.0020	0.0020	0.0023	0.0020	0.0022	0.0020
$std(R, \bar{\theta}_i)$	0.0028	0.0027	0.0031	0.0028	0.0031	0.0028

In the next section we begin the analysis of optimal policies in and across models.

4.4.2 Optimal policy at the posterior mean

To establish a standard and to explain the stabilization trade-off, we determine the optimal policy $\phi_i^\star = (\rho_R^\star, \phi_\pi^\star, \phi_y^\star)_i$ at the posterior mean $\bar{\theta}_i$ for each Model i, $i = 1, 2, .., \text{C}$. To ease the numerical computation and to exclude unreasonably high policy responses, we assume the following bounds for the policy coefficients of the simple interest rate rule:

$$\rho_R \in [0, 20], \quad \rho_\pi \in [0, 20], \text{ and } \rho_y \in [0, 20].$$

The optimal coefficients and the resulting business cycles costs ($\mathcal{BC}$) expressed as equivalent reductions in steady-state consumption are displayed in Table 4.3.

Table 4.3: Optimal policy at the posterior mean ($\phi_i^\star$)

$$\widehat{R}_t = \rho_R \widehat{R}_{t-1} + \phi_\pi \widehat{\pi}_t + \phi_y \widehat{y}_t$$

	M_1	M_2	M_3	M_4	M_C
ρ_R	0.81	1.05	0.62	1.26	1.36
ϕ_π	20.00	20.00	20.00	2.42	1.01
ϕ_y	0.00	0.02	0.00	0.00	0.00
$\mathcal{BC}(\bar{\theta}_i, \phi_i^\star)$	0.0014%	0.0014%	0.0020%	0.0194%	0.0178%

Optimal policies are characterized by drawing on past interest rates. Put differently, optimal policy is history-dependent (Woodford, 2003a,b). In the first three models inflation stabilization is the predominant aim. Correspondingly, optimal policies feature a strong reaction on inflation.[10] In Models 4 and C, households value real money balances as a medium for transactions. This introduces stabilization of the nominal interest rate as a conflicting aim to price stability (see (4.25) and (4.27)) in the presence of fluctuations in the natural rate of interest. For intuition on this, suppose that ϕ_y is small and that the economy in Model 1 is hit by a wage-markup shock. To fight inflationary tendencies the output gap must decrease according to the aggregate supply curve (4.16). This in turn requires a strong increase in the nominal interest rate to fulfil the Euler equation (4.15), since the cost-push shock affects the natural rate of interest. Therefore, optimal policies in models with a demand for cash exhibit a higher coefficient ρ_R to smooth interest rates and a less aggressive response to inflation.

[10]The optimal policy response on inflation in these models always corresponds to its upper bound. However, the welfare comparison between the two approaches to model uncertainty is independent of the particular choice of the upper bound on the inflation response.

Welfare costs in models that feature a transaction friction are substantially higher. This increase is due to two effects. First, the stabilization of the interest rate adds a new component to the welfare-relevant stabilization loss, which accounts for over fifty percent of the increase in business cycle costs in Model 4 relative to Model 1. The second effect relates to the conflict of stabilizing interest rates, inflation and the output gap simultaneously, as apparent in the muted response to inflation in the optimal rules for Models 4 and C. The resulting increase in the unconditional weighted variances of inflation and the output gap accounts for the remaining increase in the costs of business cycle fluctuations.

Table 4.4: The weights λ_d and λ_R at the posterior mean

Weights	M_1	M_2	M_3	M_4	M_{C}
λ_d	0.0063	0.0231	0.0079	0.0057	0.0328
λ_R	-	-	-	0.0602	0.0728

Table 4.4 shows how the importance of stabilization aims relative to inflation for households changes across models. For example, the stabilization of the output gap is five times more important in Model C than in Model 1. In addition, the exact gap that policy should stabilize to maximize welfare differs (see (4.18) and (4.27)). Furthermore, comparing the two models that feature a demand for cash reveals that the optimal response to changes in inflation is larger in Model 4 than in Model C. Although both specifications incorporate stabilizing the nominal interest rate as a policy aim, this aim is relatively more important in Model C than in Model 4.

4.4.3 Evaluating two approaches to model uncertainty

In this section we quantitatively compare the two approaches to model uncertainty, the complete-model and the model-averaging approach. We start by determining the set of policy coefficients for the former approach according to (4.4), which yields

$$\phi^{\star}_{\text{C}}: \quad \rho_R = 1.34; \quad \phi_\pi = 1.17; \quad \phi_y = 0.00.$$

However, Model C is not the likeliest model since it also contains features which do not help to explain the given time series of GDP, inflation and the real wage (see Table 4.1). A policymaker pursuing a model-averaging approach to model uncertainty weights welfare losses in a particular model with its posterior probability, i.e. derives an optimal policy over all draws and models according to (4.7):

$$\phi_a^\star : \quad \rho_R = 1.39; \quad \phi_\pi = 3.36; \quad \phi_y = 0.00.$$

Comparing the characteristics of the two rules reveals two similarities and one difference. Both rules draw heavily on past interest rates to avoid welfare-reducing fluctuations in the interest rate in Models 4 and C, and put no emphasize on stabilizing the output gap. The main difference between both rules is the preference to stabilize inflation. While there is a conflict in stabilizing inflation and the nominal interest rate jointly in Model C, this trade-off is absent in the likeliest model, Model 3.

To evaluate the performance of the two approaches as a guard against model uncertainty we compute the business cycle cost for both policy rules in each Model i, i.e. $\mathcal{BC}(\Theta_i, \phi_{\text{C}}^\star)$ and $\mathcal{BC}(\Theta_i, \phi_a^\star)$ for $i = 1, 2, 3, 4, \text{C}$.

Table 4.5: Relative performance of $\phi_{\text{C}}^\star$ and $\phi_a^\star$

	M_1	M_2	M_3	M_4	M_{C}	WA
$\mathcal{BC}(\Theta_i, \phi_{\text{C}}^\star)/\mathcal{BC}(\Theta_i, \phi_a^\star)$	2.22	2.16	1.90	1.26	0.74	1.68

WA denotes the posterior-model probability average of business cycle costs.

As can be seen from Table 4.5, the optimal rule $\phi_a^\star$ performs twice as good as $\phi_{\text{C}}^\star$ in Models 1, 2 and 3 where inflation stabilization is the predominant principle. Nevertheless, by reacting less harshly to inflation than the optimal rules from those models (see Table 4.3), it avoids high welfare losses in Model C. On average, optimal policy derived from the model-averaging approach leads to welfare gains of 68% relative to the optimal policy rule derived by the complete-model approach.

4.5 Conclusion

In this chapter we have analyzed how to optimally conduct policy from a Bayesian perspective when the policymaker faces uncertainty about the appropriate choice among nested models. In particular, we have compared two approaches to model uncertainty. The complete-model approach is set to represent the standard practice: without paying special attention to the set of sub-models, the policymaker determines optimal policy by maximizing households' utility within one single model that nests all features and frictions. The model-averaging approach takes uncertainty about the appropriate choice of nested models into account and weights over the set of nested models to derive optimal policy prescriptions. Using EU-13 data, we find that the model-averaging approach leads to welfare gains of approximately 70 percent compared to the standard practice.

Appendix A

Technical Appendix to chapter 2

A.1 Derivation of the posterior distribution of the BVAR

A.1.1 Prior distribution

$$vec(B)|\Sigma \sim \mathcal{N}(vec(B_0), \Sigma \otimes N_0^{-1}) \tag{A.1}$$

$$\Sigma \sim \mathcal{IW}(v_0 S_0, v_0) \tag{A.2}$$

Σ is of size $m \times m$, N_0 of size $k \times k$, where $k = m * l$. The probability density function (p.d.f.) of $vec(B)$ is given by:

$$\begin{aligned} p(B|B_0, \Sigma, N_0) = & (2\pi)^{-mk/2} |\Sigma \otimes N_0^{-1}|^{-1/2} \\ & exp\left[-\frac{1}{2}\left(vec(B) - vec(B_0)\right)'\left(\Sigma^{-1} \otimes N_0\right)\left(vec(B) - vec(B_0)\right)\right] \\ & = (2\pi)^{-mk/2} |\Sigma|^{-k/2} |N_0|^{m/2} exp\{-\frac{1}{2} tr\left[\Sigma^{-1}\left(B - B_0\right)' N_0 \left(B - B_0\right)\right]\}. \end{aligned}$$

The p.d.f. of Σ is defined as:

$$p(\Sigma|v_0 S_0, v_0) = C_{IW}^{-1} |\Sigma|^{-\frac{1}{2}(v_0+m+1)} exp\left[-\frac{1}{2} tr\left(\Sigma^{-1} v_0 S_0\right)\right],$$

where:

$$C_{IW} = 2^{\frac{1}{2} v_0 m} \pi^{\frac{1}{4} m(m-1)} \prod_{i=0}^{m} \Gamma\left(\frac{v_0 + 1 - i}{2}\right) |S_0|^{-\frac{1}{2} v_0}.$$

A.1.2 Likelihood

For

$$vec(u) \sim \mathcal{N}(0, \Sigma \otimes I), \tag{A.3}$$

$$p(Y|B,\Sigma) = (2\pi)^{-Tm/2}|\Sigma|^{-T/2} exp\{-\frac{1}{2} tr \left[\Sigma^{-1}(Y-XB)'(Y-XB)\right]\}. \tag{A.4}$$

The kernel can be rewritten as:

$$\begin{aligned}(Y-XB)'(Y-XB) &= (Y-XB-X\hat{B}+X\hat{B})'(Y-XB-X\hat{B}+X\hat{B}) \\ &(Y-X\hat{B})'(Y-X\hat{B}) + (B-\hat{B})'X'X(B-\hat{B}).\end{aligned} \tag{A.5}$$

A.1.3 Posterior

$$\begin{aligned}p(\Sigma, B|Y) = {} & C_{IW}^{-1}|\Sigma|^{-\frac{1}{2}(v_0+m+1)} exp\left[-\frac{1}{2} tr\left(\Sigma^{-1} v_0 S_0\right)\right] \\ & \times (2\pi)^{-mk/2}|\Sigma|^{-k/2}|N_0|^{m/2} exp\{-\frac{1}{2} tr\left[\Sigma^{-1}\left(B-B_0\right)' N_0 \left(B-B_0\right)\right]\} \\ & \times (2\pi)^{-Tm/2}|\Sigma|^{-T/2} exp\{-\frac{1}{2} tr\left[\Sigma^{-1}(Y-X\hat{B})'(Y-X\hat{B})\right]\} \\ & \times exp\{-\frac{1}{2} tr\left[\Sigma^{-1}(B-\hat{B})'X'X(B-\hat{B})\right]\}\end{aligned} \tag{A.6}$$

Use the formula as stated in Leamer (1978)[1]:

$$\begin{aligned}(B-\hat{B})'X'X(B-\hat{B})\left(B-B_0\right)' N_0 \left(B-B_0\right) &= (B-B_T)'N_T(B-B_T) \\ &\times (B-B_0)'(X'X(N_T)^{-1}N_0)(B-B_0),\end{aligned} \tag{A.7}$$

where:

$$\begin{aligned}N_T &= N_0 + X'X \\ B_T &= N_T^{-1}(N_0 B_0 + X'X\hat{B})\end{aligned}$$

[1] Appendix 1, T10

leads to:

$$
\begin{aligned}
p(\Sigma, B|Y) = C_{IW}^{-1}|\Sigma|^{-\frac{1}{2}(v_0+m+1)} exp\left[-\frac{1}{2}tr\left(v_0\Sigma^{-1}S_0\right)\right] \\
\times (2\pi)^{-mk/2}|\Sigma|^{-k/2}|N_0|^{m/2} \\
\times exp\{-\frac{1}{2}tr\left[\Sigma^{-1}(B-B_0)'(X'X(N_T)^{-1}N_0)(B-B_0)\right]\} \\
\times (2\pi)^{-Tm/2}|\Sigma|^{-T/2}exp\{-\frac{1}{2}tr\left[\Sigma^{-1}(Y-X\hat{B})'(Y-X\hat{B})\right]\} \\
\times exp\{-\frac{1}{2}tr\left[\Sigma^{-1}(B-B_T)'N_T(B-B_T)\right]\}
\end{aligned}
\tag{A.8}
$$

$$
\begin{aligned}
p(\Sigma, B|Y) = C_{IW}^{-1}|\Sigma|^{-\frac{1}{2}(T+v_0+m+1)} \\
\times exp[-\frac{1}{2}tr(\Sigma^{-1}(\frac{v_0}{v_T}S_0 + \frac{T}{v_T}\tilde{\Sigma} \\
+ \frac{1}{v_T}(B-B_0)'(X'X(N_T)^{-1}N_0)(B-B_0)))] \\
\times (2\pi)^{-m(T+k)/2}|\Sigma|^{-(T+k)/2}exp\{-\frac{1}{2}tr\left[\Sigma^{-1}(B-B_T)'N_T(B-B_T)\right]\}.
\end{aligned}
\tag{A.9}
$$

A.2 Description and solution of the FTPL model

A.2.1 FTPL Model Setup

$$U_{F,t} = log(c_{F,t}) + log(m_{F,t}) \tag{A.10}$$

$$c_{F,t} + m_{F,t} + b_{F,t} + \tau_{F,t} = y_F + \frac{1}{\pi_{F,t}}m_{F,t-1} + \frac{R_{F,t-1}}{\pi_{F,t}}b_{F,t} \tag{A.11}$$

First-order conditions:

$$\frac{1}{R_{F,t}} = \beta_F\frac{1}{\pi_{F,t+1}} \tag{A.12}$$

$$m_{F,t} = c_F\frac{R_{F,t}}{R_{F,t}-1} \tag{A.13}$$

Government budget constraint:

$$b_{F,t} + m_{F,t} + \tau_{F,t} = g_F + \frac{M_{F,t-1}}{P_{F,t}} + R_{F,t-1}\frac{B_{F,t-1}}{P_{F,t}} \tag{A.14}$$

Monetary authority:

$$R_{F,t} = \alpha_{F0} + \alpha_F \pi_{F,t} + \theta_{F,t} \tag{A.15}$$
$$\theta_{F,t} = \rho_{F1}\theta_{F,t-1} + \epsilon_{F1,t} \tag{A.16}$$
$$\epsilon_{F1,t} \sim N(0, \sigma_{F1}) \tag{A.17}$$

Fiscal authority:

$$\tau_{F,t} = \gamma_{F0} + \gamma b_{F,t-1} + \psi_{F,t} \tag{A.18}$$
$$\psi_{F,t} = \rho_{F2}\psi_{F,t-1} + \epsilon_{F2,t} \tag{A.19}$$
$$\epsilon_{F2,t} \sim N(0, \sigma_F 2) \tag{A.20}$$

A.2.2 Linearization

$$\bar{x}\hat{x}_t = \tilde{x}_t.$$

First equation:

$$\begin{aligned} R_{F,t} &= \alpha_{F0} + \alpha_F \pi_{F,t} + \theta_{F,t} \\ \pi_{F,t+1} &= \beta_F \alpha_{F0} + \beta_F \alpha_F \pi_{F,t} + \beta_F \theta_{F,t} \\ \tilde{\pi}_{F,t+1} &= \beta_F \alpha_f \tilde{\pi}_{F,t} + \beta_F \theta_{F,t} \end{aligned}$$

Second equation:

$$\begin{aligned} \bar{R}_F \hat{R}_{F,t} &= \frac{\bar{\pi}_F}{\beta_F} \hat{\pi}_{F,t+1} \\ \tilde{R}_{F,t} &= \frac{\tilde{\pi}_{F,t+1}}{\beta_F} \end{aligned}$$

$$\begin{aligned} m_{F,t} &= c_F \frac{R_{F,t}}{R_{F,t} - 1} \\ \tilde{m}_{F,t} &= -\frac{c_F}{(\bar{R}_F - 1)^2 \beta_F} \tilde{\pi}_{F,t+1} \end{aligned}$$

$$\tilde{m}_{F,t} = -\frac{c_F}{(\bar{R}_F - 1)^2\beta_F}(\beta_F\alpha_F\tilde{\pi}_{F,t} + \beta_F\theta_{F,t})$$
$$\tilde{m}_{F,t-1} = -\frac{c_F\alpha_F}{(\bar{R}_F - 1)^2}\tilde{\pi}_{F,t-1} - \frac{c_F}{(\bar{R}_F - 1)^2}\theta_{F,t-1}$$

$$b_{F,t} + m_{F,t} + \tau_{F,t} = g_F + \frac{M_{F,t-1}}{P_{F,t}} + R_{F,t-1}\frac{B_{F,t-1}}{P_F, t}$$
$$\tilde{b}_{F,t} + \tilde{m}_{F,t} + \tilde{\tau}_{F,t} = \frac{\tilde{m}_{F,t-1}}{\bar{\pi}_F} - \frac{\bar{m}_F}{\bar{\pi}_F^2}\tilde{\pi}_{F,t} + \frac{\bar{b}_F}{\bar{\pi}_F}\tilde{R}_{F,t-1} - \frac{\bar{R}_F\bar{b}_F}{\bar{\pi}_F^2}\tilde{\pi}_{F,t} + \frac{\bar{R}_F}{\bar{\pi}_F}\tilde{b}_{F,t-1}$$

$$\tilde{b}_{F,t} - \frac{c_F\alpha_F}{(\bar{R}_F - 1)^2}\tilde{\pi}_{F,t} + \frac{c_F\bar{R}_F}{(\bar{R}_F - 1)\bar{\pi}_F^2}\tilde{\pi}_{F,t} + \frac{\bar{R}_F\bar{b}_F}{\bar{\pi}_F^2}\tilde{\pi}_{F,t} - \frac{c_F}{(\bar{R}_F - 1)^2}\theta_{F,t} + \tilde{\tau}_{F,t}$$
$$= \frac{\tilde{m}_{F,t-1}}{\bar{\pi}_F} + \frac{\bar{b}_F}{\bar{\pi}_F}\tilde{R}_{F,t-1} + \frac{\bar{R}_F}{\bar{\pi}_F}\tilde{b}_{F,t-1}$$

A.2.3 Simplifying the FTPL model

Define:

$$\varphi_{F1} = \frac{c_F}{(\bar{R}_F - 1)}\left(-\frac{\alpha_F}{(\bar{R}_F - 1)} + \frac{c_F}{\beta_F\bar{\pi}_F}\right) + \frac{\bar{b}_F}{\beta_F\bar{\pi}_F}$$

$$\varphi_{F3} = -\frac{c_F}{(\bar{R}_F - 1)^2}$$

$$-\frac{1}{\bar{\pi}_F}\frac{c_F\alpha_F}{(\bar{R}_F - 1)^2} + \frac{\bar{b}_F}{\bar{\pi}_F}\alpha_F = -\frac{\alpha_F}{\bar{\pi}_F}\left[\frac{c_F}{(\bar{R}_F - 1)^2} - \bar{b}_F\right] = -\varphi_{F2}$$
$$-\frac{1}{\bar{\pi}_F}\frac{c_F}{(\bar{R}_F - 1)^2} + \frac{\bar{b}_F}{\bar{\pi}_F} = -\frac{1}{\bar{\pi}_F}\left[\frac{c_F}{(\bar{R}_F - 1)^2} - \bar{b}_F\right] = \frac{-\varphi_{F2}}{\alpha_F} = -\varphi_{F4}$$

This yields:

$$\tilde{b}_{F,t} + \varphi_{F1}\tilde{\pi}_{F,t} + \varphi_{F3}\theta_{F,t} - (\beta_F^{-1} - \gamma_F)\tilde{b}_{F,t-1} + \psi_{F,t} + \varphi_{F4}\theta_{F,t-1} + \varphi_{F2}\tilde{\pi}_{F,t-1} = 0 \quad \text{(A.21)}$$

$$\tilde{b}_{F,t} + \varphi_{F1}\tilde{\pi}_{F,t} + \varphi_{F3}\theta_{F,t} + \psi_{F,t} = (\beta_F^{-1} - \gamma_F)\tilde{b}_{F,t-1} - \varphi_{F4}\theta_{F,t-1} - \varphi_{F2}\tilde{\pi}_{F,t-1} \quad \text{(A.22)}$$

A.2.4 Calibration

Following Leeper (1991) the FTPL model is calibrated by setting:

$$\begin{aligned}
\beta_F &= 0.99 \\
\bar{c}_F &= 0.75 \\
\frac{\bar{b}_F}{\bar{y}_F} &= 0.4 \\
\bar{\pi}_F &= 3.43 \\
\rho_{F1} &= 0.8 \\
\rho_{F2} &= 0 \\
\sigma_{F1} &= 0.2 \\
\sigma_{F2} &= 0.2
\end{aligned}$$

A.3 Solution of the Deep habits model

A.3.1 Steady state

$$R^{\star}_H = 1/\beta_H \tag{A.23}$$
$$\bar{h}_H = 0.3 \tag{A.24}$$
$$\bar{x}_H = (1-\theta^d_H)\bar{h}_H \tag{A.25}$$
$$\bar{c}_H = \bar{h}_H \tag{A.26}$$
$$\bar{\lambda}^c_H = 1/((1-\theta^d_H)\eta_H) \tag{A.27}$$
$$\bar{\lambda}^y_H = 1 + (\theta^d_H\beta_H - 1)\bar{\lambda}^c_H \tag{A.28}$$
$$\bar{w}_H = \qquad \bar{\lambda}^y_H \tag{A.29}$$
$$\bar{\lambda}^h_H = \bar{w}_H/(\bar{x}^{\sigma}_H\phi_H) \tag{A.30}$$
$$\gamma_H = (\bar{x}^{-\sigma}_H\bar{w}_H - \bar{\lambda}^{\bar{h}_H}_H)/\bar{h}^{\kappa}_H \tag{A.31}$$

A.3.2 Loglinearized equations

$$\bar{x}_H\hat{x}_{H,t} = \bar{c}_H\hat{c}_{H,t} - \theta^d_H\bar{c}_H\hat{c}_{H,t-1} \tag{A.32}$$

$$\gamma_H\bar{h}^{\kappa_H}_H\kappa_H\hat{h}_{H,t} = \bar{x}^{-\sigma_H}_H\bar{w}_H(-\sigma_H\hat{x}_{H,t} + \hat{w}_{H,t}) - \bar{\lambda}_H\hat{\lambda}^h_{H,t} \tag{A.33}$$

$$\begin{aligned}\phi_H\bar{\lambda}^h_H\bar{h}_H\bar{x}^{\sigma}_H(\hat{\lambda}^h_{H,t} + \hat{h}_{H,t} + \sigma_H\hat{x}_{H,t}) + \zeta_{Hw}(\hat{\pi}_{Hw,t} - \hat{\bar{\pi}}_{Hw,t})\\ = \bar{h}_H\bar{w}_H(\hat{h}_{H,t} + \hat{w}_{H,t}) + \beta\zeta_{Hw}(\hat{\pi}_{Hw,t+1} - \hat{\bar{\pi}}_{Hw,t+1})\end{aligned} \tag{A.34}$$

$$-\sigma_H\hat{x}_{H,t} = R^{\star}_H\hat{R}_{H,t} - \sigma\hat{x}_{H,t+1} - \hat{\pi}_{H,t+1} \tag{A.35}$$

$$\hat{c}_{H,t} = \hat{h}_{H,t} \tag{A.36}$$
$$\hat{\lambda}^y_{H,t} = \hat{w}_{H,t} \tag{A.37}$$
$$\hat{h}_{H,t} = \hat{y}_{H,t} \tag{A.38}$$

$$\bar{\lambda}_H^y \hat{\lambda}_{H,t}^y + \bar{\lambda}_H^c \hat{\lambda}_{H,t}^c = \theta_H^d \beta_H \bar{\lambda}_H^c (-\sigma_H \hat{x}_{H,t+1} + \sigma_H \hat{x}_{H,t} + \hat{\lambda}_{H,t+1}^c) \tag{A.39}$$

$$\eta_H \bar{\lambda}_H^c \bar{x}_H (\hat{\lambda}_{H,t}^c + \hat{x}_{H,t}) + \zeta_{Hp}(\hat{\pi}_{H,t} - \hat{\bar{\pi}}_{H,t}) = \bar{c}_H \hat{c}_{H,t} + \beta_H \zeta_{Hp}(\hat{\pi}_{H,t+1} - \hat{\bar{\pi}}_{H,t+1}) \tag{A.40}$$

$$\hat{R}_{H,t} = \rho_{Hr} \hat{R}_{H,t-1} + (1-\rho_{Hr})(\alpha_{H\pi}\hat{\pi}_{H,t} + \alpha_{Hy}\hat{y}_{H,t}) + \epsilon_{H,t} \tag{A.41}$$

$$\hat{\bar{\pi}}_{H,t} = (1-\nu_{Hp})\hat{\pi}_{H,t-1} \tag{A.42}$$

$$\hat{\bar{\pi}}_{Hw,t} = (1-\nu_{Hw})\hat{\pi}_{Hw,t-1} \tag{A.43}$$

$$\bar{w}_H \hat{w}_{H,t} = \bar{w}_H \hat{w}_{H,t-1} + \hat{\pi}_{Hw,t} - \hat{\pi}_{H,t} \tag{A.44}$$

A.3.3 Calibration and estimation of the deep habits model

Following Ravn et al. (2008), calibrated values for the structural parameters are set as:

$$\begin{aligned}
R_H^\star &= 1.01 \\
\beta_H &= 1/R_H^\star \\
\phi_H &= 4 \\
\kappa_H &= 0.5 \\
\pi_H^\star &= 1 \\
\sigma_H &= 3 \\
\pi_{Hw}^\star &= 1 \\
\nu_{Hw} &= 0.96
\end{aligned}$$

A.4 Data description

The frequency of all data used is quarterly. The data ranges from 1955.1 to 2009.1. All series except the Fed Funds rate are in logs. GDP, personal consumption and real wages are transformed into per capita.

Nominal GDP: This series is *BEA NIPA Table 1.1.5. Gross Domestic Product.*

Private Consumption: This series is *BEA NIPA Table 1.1.5. Personal consumption expenditures.*

Wage: The wage rate is the series COMPNFB (Nonfarm Business Sector: Compensation Per Hour) at the Federal Reserve Board of St. Louis' website.

Interest Rate: This is the Federal Funds rate taken from Fred2.

Adjusted reserves: This is the adjusted monetary base given by the series *adjressl http://research.stlouisfed.org/fred2/series/ADJRESSL.*

PPIC: This series is *http://research.stlouisfed.org/fred2/series/PPICRM.*

Real GDP: This series is *BEA NIPA Table 1.1.6. Real Gross Domestic Product.*

Implicit GDP Deflator: The implicit GDP deflator is calculated as the ratio of **N**ominal GDP to **R**eal GDP

Civilian Population: This is a quarterly measure for the population given by the respective average of the monthly values of the series *CNP16OV, Civilian Noninstitutional Population* at the Federal Reserve Board of St. Louis' website *http://research.stlouisfed.org/fred2/.* The numbers have been converted from thousands to billions.

Table A.1: Prior distribution, Monte Carlo results and Posterior estimates of the structural parameters for the Deep habits model. Columns 1-4 specify the name and type of prior distribution with corresponding mean and standard deviation. Column 5 displays the parameter value for the data that was simulated from the DSGE model. Columns $mean_1$, std_1 give the Monte Carlo estimation results for the small parametrization of the DSGE model, columns $mean_2$, std_2 for the full parametrization. The last two columns display the estimation results from confronting the DSGE model with the data.

		Prior distribution		*Monte Carlo Experiment*					*Posterior distribution*	
Parameter	distribution	mean	std	sim	$mean_1$	std_1	$mean_2$	std_2	mean	std
θ_H^d	beta	0.5	0.2	0.85	0.84	0.01	0.84	0.01	0.72	0.12
η_H	normal	2	0.3	2.5	2.44	0.12	2.11	0.11	2.47	0.60
ζ_{Hw}	normal	40	5	40.9	40.21	1.23	40.60	0.55	42.50	15.87
ζ_{Hp}	normal	7	3	14.5	13.71	1.51	13.50	1.75	14.89	5.08
ν_{Hp}	normal	0.1	0.01	0	-	-	0.10	0.01	0.10	0.01
ρ_{Hr}	beta	0.5	0.2	0.7	-	-	0.67	0.03	0.81	0.07
$\alpha_{H\pi}$	beta	1.5	0.25	1.46	-	-	1.40	0.12	1.56	0.68
α_{Hy}	beta	0.125	0.1	0.1	-	-	0.05	0.02	0.01	0.25
σ_{Hr}	normal	2	8 *d.o.f.*	1	0.6	0.15	0.64	0.13	0.36	0.10

A.5 Figures

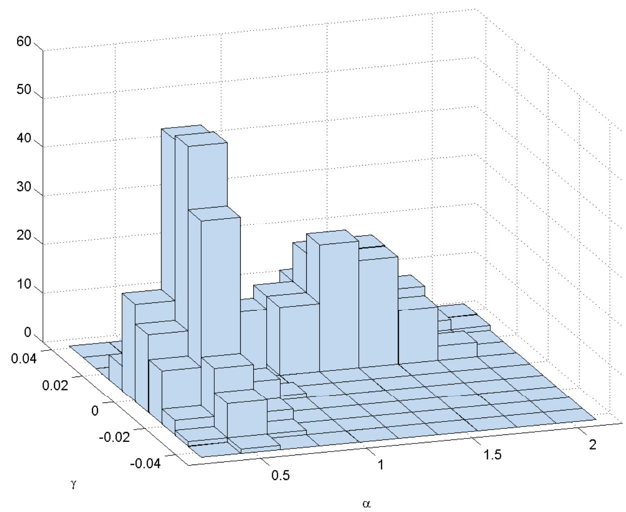


Figure A.1: Prior distribution FTPL model

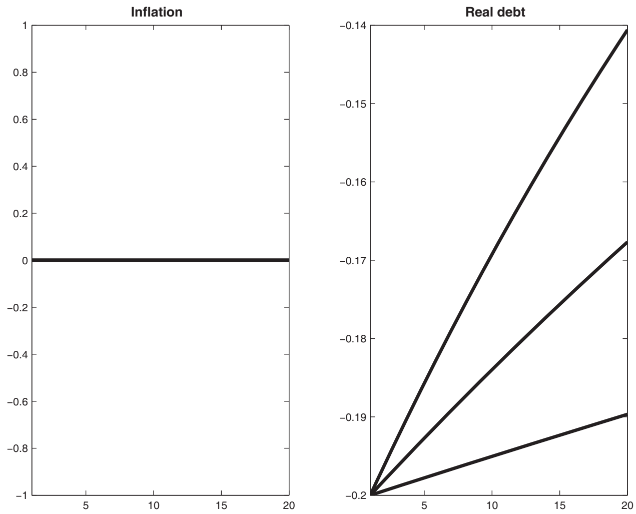


Figure A.2: Prior Bayesian IRF for a fiscal policy shock regime I in the FTPL model

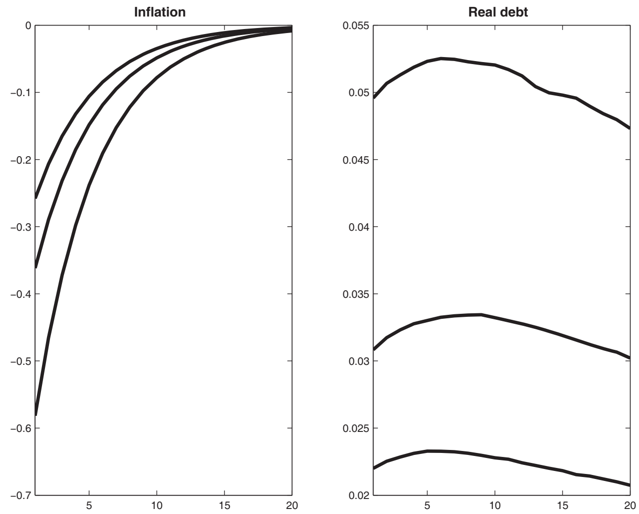


Figure A.3: Prior Bayesian IRF for a monetary policy shock regime I in the FTPL model

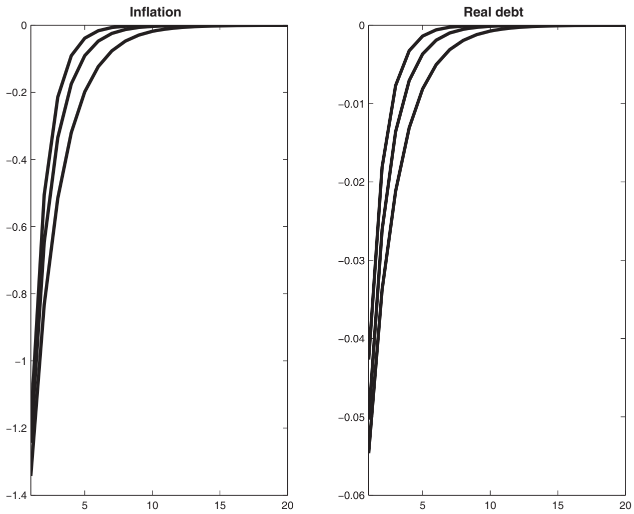


Figure A.4: Prior Bayesian IRF for a fiscal policy shock regime II in the FTPL model

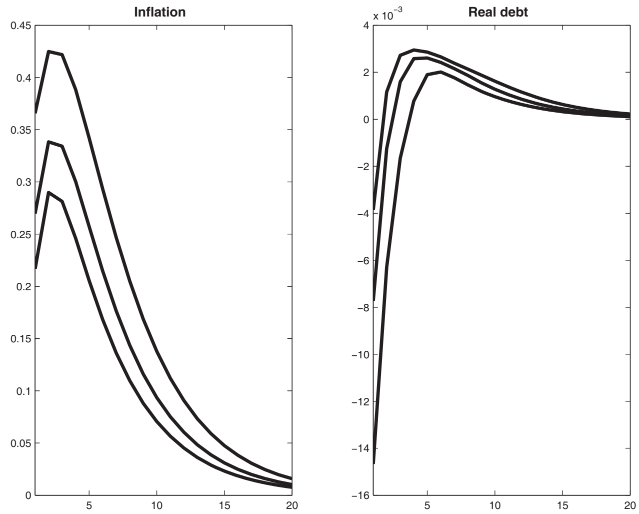


Figure A.5: Prior Bayesian IRF for a monetary policy shock regime II in the FTPL model

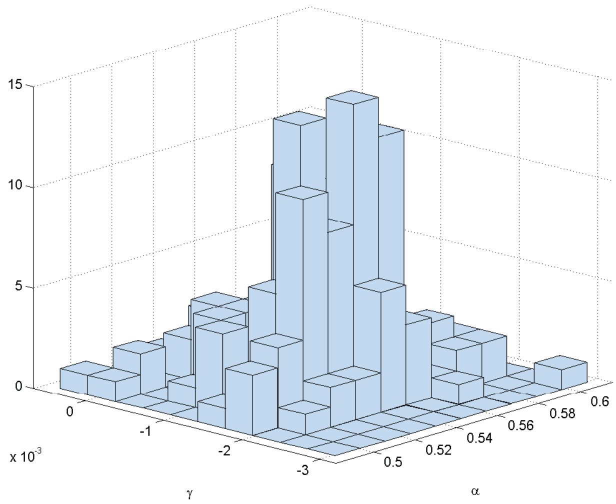


Figure A.6: Posterior distribution FTPL model

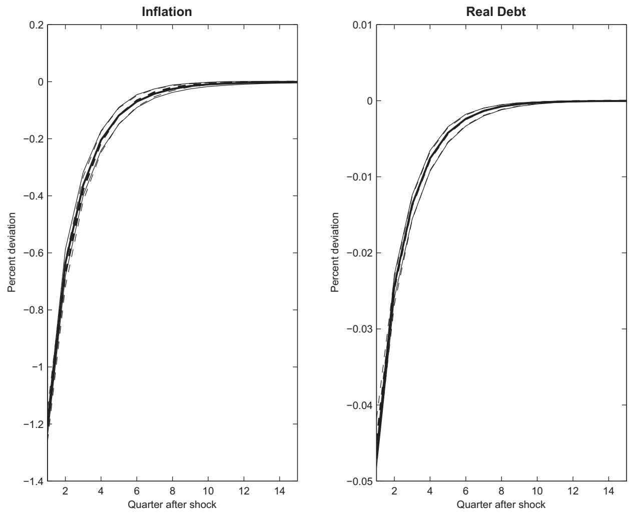


Figure A.7: Estimated Bayesian IRF for a fiscal policy shock in the FTPL model: VAR model (black line) vs. DSGE model (dashed line).

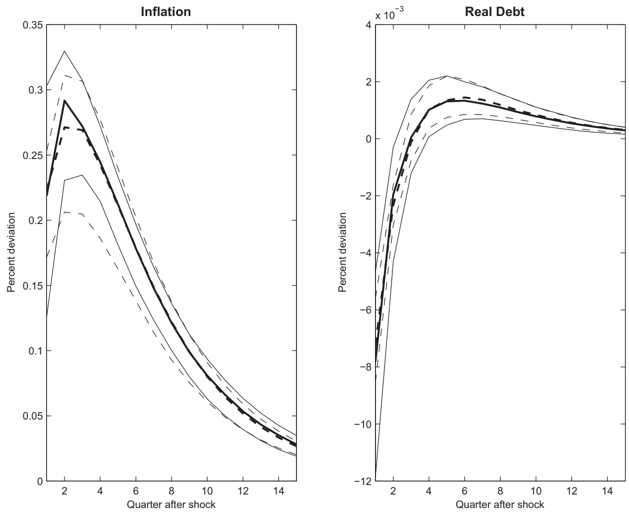


Figure A.8: Estimated Bayesian IRF for a monetary policy shock in the FTPL model: VAR model (black line) vs. DSGE model (dashed line).

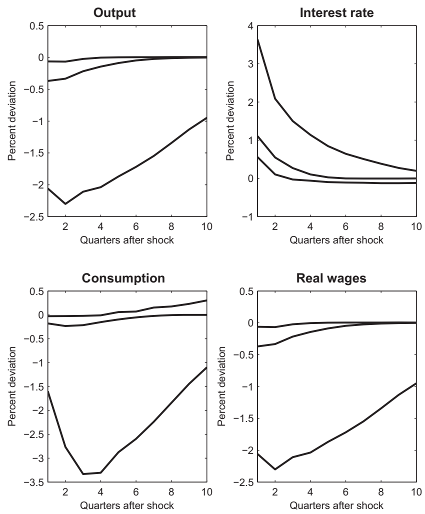


Figure A.9: Impulse response function of the deep habits model drawing from the prior distribution of deep parameters (100 % probability bands).

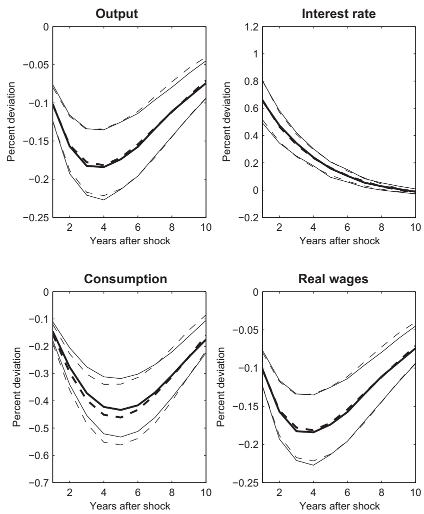


Figure A.10: Impulse response functions of the deep habits model (dashed line) versus VAR model with simulated data (68 % probability bands).

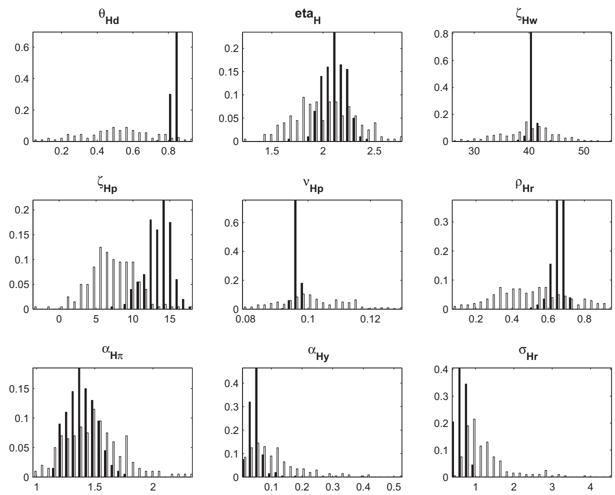


Figure A.11: Prior distribution (white) vs. Posterior distribution (black). Monte-Carlo experiment deep habits model.

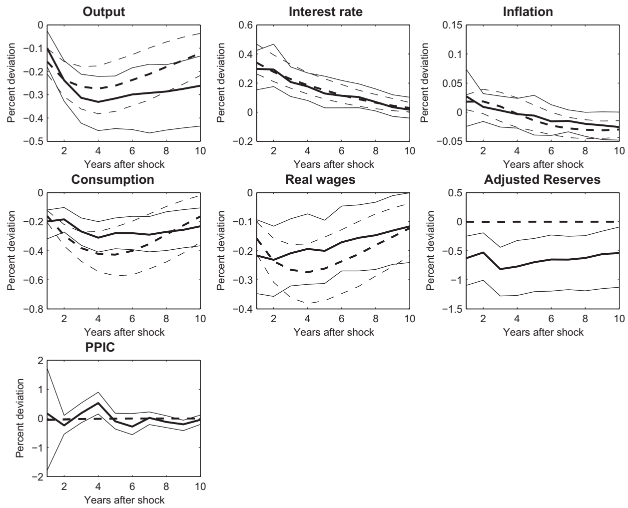


Figure A.12: Posterior distribution of impulse response functions of the deep habits model (dashed line) versus VAR model (solid line) (68 % probability bands).

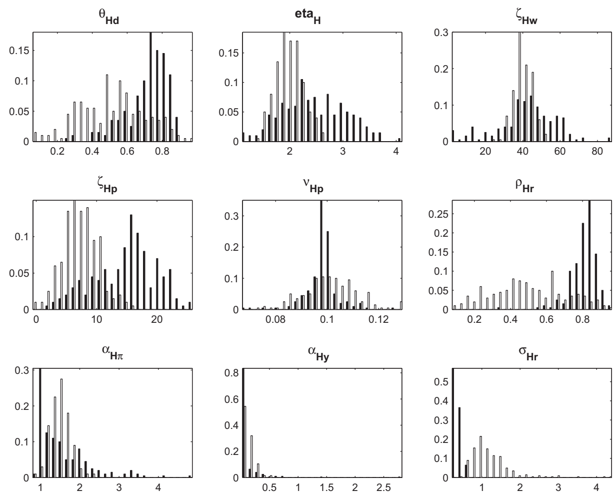


Figure A.13: Prior distribution (white) vs. Posterior distribution (black). deep habits model.

Appendix B

Technical Appendix to chapter 3

B.1 Data description

The frequency of all data used is quarterly.

Real GDP: This series is *BEA NIPA table 1.1.6 line 1 (A191RX1).*

Nominal GDP: This is a measure for the nominal GDP given by the series *GDP, Gross Domestic Product* at the Federal Reserve Board of St. Louis' website *http://research.stlouisfed.org/fred2/*. It is measured in billions of dollars.

Implicit GDP Deflator: The implicit GDP deflator is calculated as the ratio of **N**ominal GDP to **R**eal GDP.

Private Consumption: Nominal consumption expenditures for non-durables and services is the sum of the respective values of the series *PCND, Personal Consumption Expenditures: Nondurable Goods* and *PCESV, Personal Consumption Expenditures: Services* at the Federal Reserve Board of St. Louis' website *http://research.stlouisfed.org/fred2/*. Both series are measured in billions of dollars.

Private Investment: Total real private investment is the sum of the respective values of the series *BEA NIPA table 1.1.6 line 6 (A006RX1)* and *PCDG, Personal Consumption Expenditures: Durable Goods* at the Federal Reserve Board of St. Louis' website *http://research.stlouisfed.org/fred2/* (billions of dollars) in real terms.

Government Expenditure: Current government expenditure is the series in *BEA NIPA table 3.1 line 15 (W022RC1).*

Government Debt: The annual government debt is the historical series that can be copied from TreasuryDirect at *http://www.treasurydirect.gov/govt/reports/pd/histdebt/histdebt.htm*. The quarterly data is generated by a linear interpolation. The values have been converted from dollars to billions of dollars. Note that the beginning of the fiscal year changed from July 1 to October 1 in 1977.

Hours worked: This series is downloadable from the website of the Bureau of Labor Statistics at *http://data.bls.gov/cgi-bin/srgate*. The series' identification number is: *PRS84006033*. It is an index (1992=100).

Wage: The wage rate is the series *COMPNFB, Nonfarm Business Sector: Compensation Per Hour* at the Federal Reserve Board of St. Louis' website *http://research.stlouisfed.org/fred2/*.

Government Consumption Expenditures: Government consumption expenditures is the series *BEA NIPA table 3.1 line 16 (A955RC1)*.

Civilian Population: This is a quarterly measure for the population given by the respective average of the monthly values of the series *CNP16OV, Civilian Noninstitutional Population* at the Federal Reserve Board of St. Louis' website *http://research.stlouisfed.org/fred2/*. The numbers have been converted from thousands to billions.

B.2 Estimation of the VMA model

B.2.1 Notation of the VMA model

Let Y_t be a $m \times 1$ vector at date $t = 1 + l, \ldots, T$, φ_i^V coefficient matrices of size $m \times m$ and ϵ an *i.i.d.* one-step ahead forecast error, distributed: $\epsilon \sim \mathcal{N}(0, I_{m \times m})$. The VMA model containing m variables is then defined by:

$$Y_t = \sum_{i=0}^{\infty} \varphi_i^V \epsilon_{t-i} \tag{B.1}$$

where φ_i^V denotes a moving average coefficient matrix. The impulse response function of a VMA model to an innovation in variable j at horizon i φ_{ij}^V is given by the j-th column of the i-th coefficient matrix.

Due to the assumption that $\Sigma_\epsilon = I_{m\times m}$ this structural moving average representation cannot be estimated directly. Instead the reduced form moving average representation with error term $u_t = A\epsilon_t$, where $u \sim \mathcal{N}(0,\Sigma)$, is estimated. The reduced form moving average coefficients are defined as $\Phi_i = \varphi_i^V A^{-1}$:

$$Y_t = \sum_{i=0}^{\infty} \Phi_i u_{t-i} \tag{B.2}$$

The factorization $\Sigma = A'A$ does not have a unique solution, which leads to an identification problem of A.

The impulse matrix $\breve{A}$ is defined as a sub matrix of A of size $m \times n$ where n is the number of structural shocks under consideration, i.e. the structural shock of interest as well as other shocks necessary to distinguish this shock. These shocks have to be included into the DSGE model as well. In order to indicate that the restrictions put on A rely on the model and therefore its parameter vector θ, I write $\breve{A}(\theta)$. Given a number of rowvectors q_j forming an orthonormal matrix Q and the lower Cholesky decomposition of Σ, $\tilde{A}$, $\breve{A}(\theta)$ is defined as: $\breve{A}(\theta) = \tilde{A}Q(\theta)$.

B.2.2 The conditional distribution of VMA model parameters

The posterior distribution is evaluated in the following way: Denote the likelihood estimates as $vec(\tilde{\Phi})$ and $vec(\tilde{\Sigma})$[1]. $vec(\Phi)$ and $vec(\Sigma)$ are normally distributed with

$$[vec(\Sigma)'vec(\Phi)']' \sim \mathcal{N}([vec(\tilde{\Sigma})'vec(\tilde{\Phi})']', \Sigma_l) \tag{B.3}$$

where Σ_l is the inverse of the Hessian computed at $[vec(\tilde{\Sigma})'vec(\tilde{\Phi})]'$. Every realization of the vector of parameters of the DSGE model θ is associated with an impulse response function of the DSGE model and a realization of the impulse matrix $\breve{A}(\theta)$.

A sequence of realizations of θ yields a sequence of restrictions and therefore a related prior probability distribution. Given a realization of an impulse response function of the DSGE model φ^D the posterior distribution is evaluated the following way:

1. Derive the sign restrictions from φ^D.

2. Draw a realization of Φ and Σ from the distribution (B.3).

3. Calculate $\tilde{A}$ and draw $Q(\theta)$ from a uniform distribution such that $\breve{A}(\theta) = \tilde{A}Q(\theta)$ fulfills the sign restriction.

[1] $vec(\Sigma)$ summarizes only the unique entries in Σ.

4. Given A compute the structural impulse responses from $\varphi_i = \Phi_i A, i = 1 \cdots k$.

B.2.3 Sampling algorithm

In order to evaluate the joint posterior distribution of the parameters of the DSGE model and the VAR model I use a Gibbs sampling algorithm combined with a Metropolis-Hastings step. The Gibbs sampling algorithm allows to draw from from the conditional distribution of the VMA model given restrictions from the DSGE model and from the conditional distribution of the DSGE model given impulse response functions from the VMA model. The Metropolis-Hastings step is a acceptance/rejection sampling algorithm that determines the probability space where the implied impulse response functions of the DSGE model and those of the VAR model coincide. It is carried out 20 times.

To initialize the algorithm candidate distributions for the conditional distributions $p(\varphi^V|\theta, Y)$ and $p(\theta|\varphi^V)$ are derived: for $j_S = 1 \ldots d_S$

1. Draw a θ_{j_S} from $p(\theta)$.

2. For every realization of the vector of deep parameters of the DSGE model derive the corresponding sign restriction.

3. Draw a Φ_{j_S} from $vec(\Phi) \sim \mathcal{N}(vec(\tilde{\Phi}), \tilde{\Sigma}_\Phi)$ and an orthonormal matrix Q_{j_S} so that $\varphi^V_{j_S}$ satisfies the restrictions derived from $\varphi^D_{j_S}$.

4. For every realization of $\varphi^V_{j_S}$ derived from step 3 find the θ_{j_S} that maximizes (2.49) combined with the prior $p(\theta)$.

5. Repeat this d times.

The d_S vectors of deep structural parameters define the candidate distribution $p(\theta|\varphi^V)^c$ and a corresponding distribution of $p(\varphi^V|\theta)^c$ for the following algorithm. At each iteration $i_S = 1, \ldots, I_S$ conduct the following steps:

1. Draw n_S times from $p(\theta|\varphi^V)^c$.

2. For every realization of the vector of deep parameters of the DSGE model derive the corresponding sign restrictions.

3. For every derived sign restrictions draw a Φ from $vec(\Phi) \sim \mathcal{N}(vec(\tilde{\Phi}), \tilde{\Sigma}_\Phi)$. Compute the lower Cholesky decomposition and find an $\breve{A} = \tilde{A}Q$ fulfilling the sign restrictions 2. Compute the corresponding φ^V, yielding $p(\varphi^V|\theta)^{\tilde{i}_S}$.

4. For every realization of φ^V derived from step 3 find the θ that maximizes (2.49) combined with the prior $p(\theta)$. This yields $p(\theta|\varphi^V)^{\tilde{i}_S}$.

5. Do acceptance-rejection by comparing $p(\theta|\varphi^V)^{\tilde{i}_S}$ with $p(\theta|\varphi^V)^{i_S-1}$. Keep the corresponding vectors from $p(\varphi^V|\theta)^{\tilde{i}_S}$. This yields $p(\theta|\varphi^V)^{i_S-1}$ and $p(\varphi^V|\theta)^{i_S}$.

6. Start again at 1.

The chain converges if $p(\theta|\varphi^V)^{i_S}$ and $p(\theta|\varphi^V)^{i_S-1}$ as well as $p(\varphi^V|\theta)^{i_S}$ and $p(\varphi^V|\theta)^{i_S-1}$ are similar, i.e. the acceptance rate is low. It is important to note that the candidate $p(\theta|\varphi^V)^c$ is not adjusting over the algorithm in order to avoid that the algorithm is stuck and to allow for a continuous wide range of the parameters of the DSGE model. Thus, d_S should be chosen high enough. I have $d_S = 200$ chosen.

B.3 Kalman Filter and root-flipping

This section provides insights how the Kalman Filter flips the roots of a non-stable process similar to a Blaschke factor and therefore recovers the correct econometric estimates. The calculations and the example in this section are mostly taken from slides by Eric Leeper[2].

To keep this section self-explanatory, I first set out briefly the DSGE model used as an example. The DSGE model, as in Leeper et al. (2008), is a standard growth model with log preferences, inelastic labor supply, and complete depreciation of capital. The equilibrium equations for consumption c_g, output v_g and capital k_g are given by:

$$\begin{aligned}
\frac{1}{c_{g,t}} &= \alpha_g \beta_g E_t (1 - \tau_{g,t+1}) \frac{1}{c_{g,t+1}} \frac{v_{g,t+1}}{k_{g,t}} \\
c_{g,t} + k_{g,t} &= v_{g,t} \\
v_{g,t} &= a_{g,t} k_{g,t-1}^{\alpha_g}
\end{aligned}$$

where β_g is the discount factor, α_g the share of capital in the production function, a_g an exogenous technology shock. The government sets taxes according to:

$$t_{g,t} = \tau_{g,t} v_{g,t}$$

After log linearizing the system of equations, the equilibrium is characterized by a

[2]From ZEI summer school 2008.

second-order difference equation in capital, which is solved by:

$$\hat{k}_{g,t} = \alpha_g \hat{k}_{g,t-1} + \hat{a}_{g,t} - (1-\theta_g)\left(\frac{\bar{\tau}_g}{1-\bar{\tau}_g}\right)\sum_{i=0}^{\infty}\theta_g^i E_t \hat{\tau}_{g,t+i+1} \tag{B.4}$$

where $\theta_g = \alpha_g\beta_g(1-\bar{\tau}_g)$. In a DSGE model with pre-announced taxes the tax rate $\hat{\tau}_t$ is a function of news shocks $\epsilon_{g,\tau,t-q}$, where q denotes the pre-announcement horizon. Define $\kappa_g = (1-\theta_g)\left(\frac{\bar{\tau}_g}{1-\bar{\tau}_g}\right)$ For $q = 2$ (B.4) can be written as:

$$\hat{k}_{g,t} = \alpha_g \hat{k}_{g,t-1} + \hat{a}_{g,t} - \kappa_g\left(\epsilon_{g,\tau,t-1} + \theta\epsilon_{g,\tau,t}\right) \tag{B.5}$$

Equation (B.5) can be written as:

$$(1-\alpha_g L)\hat{k}_{g,t} = -\kappa_g(L+\theta_g)\epsilon_{g,\tau,t}.$$

Invertibility of this stochastic process requires $|\theta_g| > 1$, which is not case. One way to achieve invertibility is to employ a Blaschke factor [3] $(L+\theta_g)/(1+\theta_g L)$. Define a new error term $\epsilon^{\star}_{g,\tau,t} = (L+\theta_g)/(1+\theta_g L)\epsilon_{g,\tau,t}$. Then:

$$\begin{aligned}
(1-\alpha_g L)\hat{k}_{g,t} &= -\kappa(L+\theta_g)\frac{(1+\theta_g L)}{(L+\theta_g)}(L+\theta_g)/(1+\theta_g L)\epsilon_{g,\tau,t} \\
(1-\alpha_g L)\hat{k}_{g,t} &= -\kappa(1+\theta_g L)\epsilon^{\star}_{g,\tau,t} \\
(1-\alpha_g L)\hat{k}_{g,t} &= -(1-\theta)\left(\frac{\bar{\tau}_g}{1-\bar{\tau}_g}\left(\epsilon^{\star}_{g,\tau,t} + \theta_g\epsilon^{\star}_{g,\tau,t-1}\right)\right) \qquad \text{(B.6)}
\end{aligned}$$

Equation (B.6) is now an invertible stochastic process for $|\theta_g| < 1$.

Another way to achieve invertibility is to write equation (B.5) as state space system in the innovation representation:

$$\begin{aligned}
x_{g,t+1} &= A_g x_{g,t} + K_g a_{g,t} \\
y_{g,t} &= C_g x_{g,t} + a_{g,t}
\end{aligned}$$

with $x_{g,t} = -(\kappa_g\theta_g)^{-1}\hat{k}_{g,t} - \epsilon_{g,\tau,t}$, $y_{g,t} = -(\kappa_g\theta_g)^{-1}\hat{k}_{g,t}$, $a_{g,t} = \epsilon_{g,\tau,t-1}$, $A_g = \alpha_g$, $C_g = 1$ and $K_g = (\alpha_g + \theta_g^{-1})$. The condition for invertibility of the system is found by using

[3] See Lippi and Reichlin (1993), Lippi and Reichlin (1994) and Mertens and Ravn (2009) for further information.

$a_{g,t} = y_{g,t} - C_g x_{g,t}$ and rewriting the state space system as:

$$\begin{aligned} x_{g,t+1} &= (A_g - C_g K_g) x_{g,t} + K_g y_{g,t} \\ a_{g,t} &= y_{g,t} - C_g x_{g,t} \end{aligned}$$

The system is stable if the eigenvalues of $(A_g - C_g K_g)$ are inside the unit circle. Since $A_g - C_g K_g = \alpha_g - \alpha_g - \theta_g^{-1} = -\theta_g^{-1}$ and $\theta_g^{-1} > 1$, the system is not stable. But, it can be shown that there exists a K_g so that the system is stable. This involves the following assumptions concerning time invariance of the Kalman Filter:

1. The pair (A'_g, C'_g) is stabilizable. A pair (A'_g, C'_g) is stabilizable if $y'_g C_g = 0$ and $y'_g A_g = \lambda_g y'_g$ for some complex number λ_g and some complex vector y_g implies that $|\lambda_g| < 1$ or $y_g = 0$.

2. The pair (A_g, G_g) is detectable. The pair (A_g, G_g) is detectable if $G'_g y_g = 0$ and $A_g y_g = \lambda_g y_g$ for some complex number λ_g and some complex vector y_g implies that $|\lambda_g| < 1$ or $y_g = 0$.

Both assumptions are fulfilled for the process and the conditions yield $\alpha_g = \lambda_g$ or $y = 0$. The Riccati equation for the covariance matrix of the innovation $a_{g,t}$ can be solved for a time invariant solution:

$$\sigma_\infty = \frac{1 - \theta_g^2}{\theta_g^2},$$

which yields a corresponding Kalman gain:

$$K_g = \alpha_g + \theta_g$$

The condition for stability reads now $A_g - K_g C_g = -\theta_g$. Since $|\theta_g| < 1$ the process is now invertible.

B.4 Loglinearized Equations of the DSGE model

The loglinearized DSGE model consists of the following equations:

$$\hat{c}_t^o = \hat{c}_{t+1}^o - \hat{R}_t + \hat{\pi}_{t+1} \tag{B.7}$$

$$\hat{q}_t = -\hat{c}_{t+1}^o + \hat{c}_t^o + (1 - \beta(1 - \delta))\hat{R}_{t+1}^k + \beta \hat{q}_{t+1} \tag{B.8}$$

$$\hat{i}_t^o - \hat{k}_{t-1}^o = \eta \hat{q}_t \tag{B.9}$$

$$\hat{k}_t^o = (1-\delta)\hat{k}_{t-1}^o + \delta \hat{i}_t^o \tag{B.10}$$

$$\hat{c}_t^r = \frac{\bar{y}}{\bar{c}} \frac{(1-\alpha)}{\mu_p}(\hat{w}_t + \hat{n}_t) - \frac{\bar{t}^r}{\bar{c}}(\phi_b \hat{b}_{t-1} + \phi_g \hat{g}_t) \tag{B.11}$$

$$\hat{c}_t + \varphi \hat{n}_t = \hat{w}_t \tag{B.12}$$

$$\hat{c}_t = \lambda \hat{c}_t^r + (1-\lambda)\hat{c}_t^o \tag{B.13}$$

$$\hat{k}_t = \hat{k}_t^o \tag{B.14}$$

$$\hat{i}_t = \hat{i}_t^o \tag{B.15}$$

$$\hat{y}_t = \alpha \hat{k}_{t-1} + (1-\alpha)\hat{n}_t \tag{B.16}$$

$$\hat{y}_t = \frac{\bar{c}}{\bar{y}}\hat{c}_t + \frac{\bar{i}}{\bar{y}}\hat{i}_t + \frac{\bar{g}}{\bar{y}}\hat{g}_t \tag{B.17}$$

$$\hat{\pi}_t = \beta E_t[\hat{\pi}_{t+1}] - \lambda_p \hat{mc}_t \tag{B.18}$$

$$\hat{mc}_t = \hat{y}_t - \hat{n}_t - \hat{w}_t \tag{B.19}$$

$$\hat{R}_t^k = \hat{c}_t + (1+\varphi)\hat{n}_t - \hat{k}_{t-1} \tag{B.20}$$

$$\frac{\bar{b}}{\bar{y}\beta}\hat{b}_t - \frac{\bar{b}}{\bar{y}\beta}\hat{R}_t + \frac{\bar{b}}{\bar{y}}\hat{\pi}_t = (\frac{\bar{b}}{\bar{y}} - \frac{\bar{t}}{\bar{y}}\phi_b)\hat{b}_{t-1} + (\frac{\bar{g}}{\bar{y}} - \frac{\bar{t}}{\bar{y}}\phi_g)\hat{g}_t \tag{B.21}$$

$$\hat{R}_t = \phi_\pi \hat{\pi}_t \tag{B.22}$$

$$\hat{g}_t = \rho_g \hat{g}_{t-1} + \epsilon_{g,t-3} \tag{B.23}$$

B.5 Tables and Figures

Table B.1: Prior and posterior distribution of the structural parameters of the DSGE model

	Prior distribution			*Posterior distribution*	
Parameter	distribution	mean	std	mean	std
ρ_g	beta	0.8	0.1	0.74	0.02
ϕ_g	gamma	0.1	0.1	0.08	0.01
ϕ_b	gamma	0.33	0.1	0.74	0.02
λ	beta	0.4	0.1	0.32	0.02
ϑ	beta	0.7	0.15	0.73	0.03
η	normal	7	5	4.2	0.81
ν	gamma	0.2	0.1	0.14	0.01
σ_g	invgamma	0.02	0.026	0.01	0.001

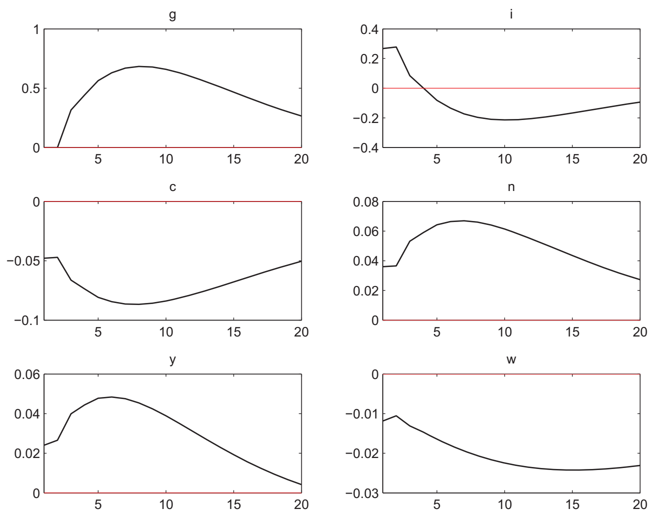


Figure B.1: Government expenditure shock two periods preannounced DSGE model calibrated to redo Ramey

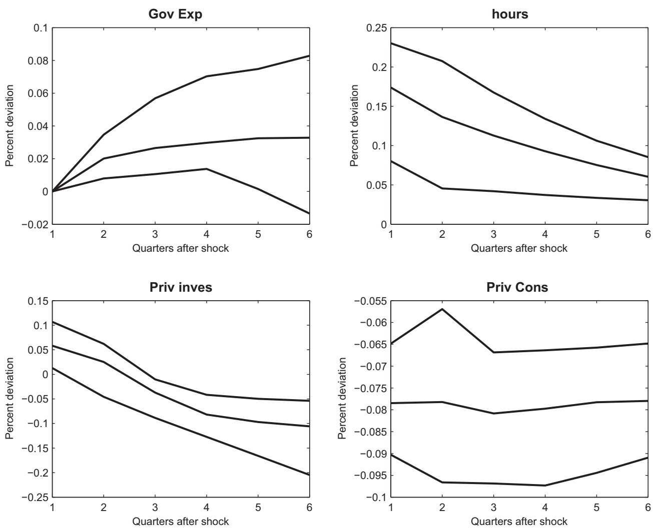


Figure B.2: Result Monte Carlo Experiment

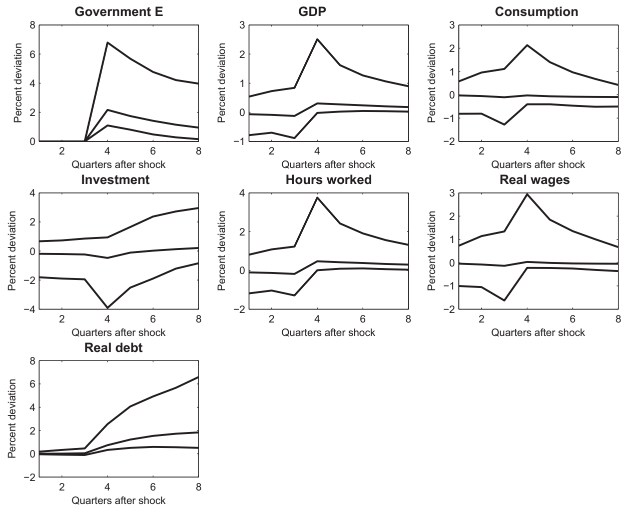


Figure B.3: Impulse response function from prior distribution of the DSGE model. All impulse responses included (100 % probability bands)

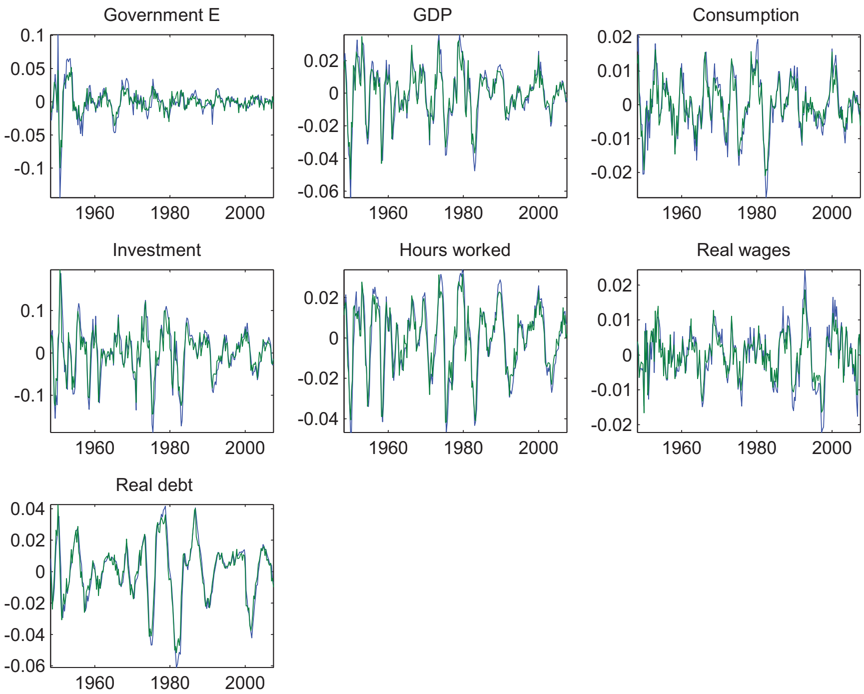


Figure B.4: Data (blue) vs. likelihood estimates VMA (green)

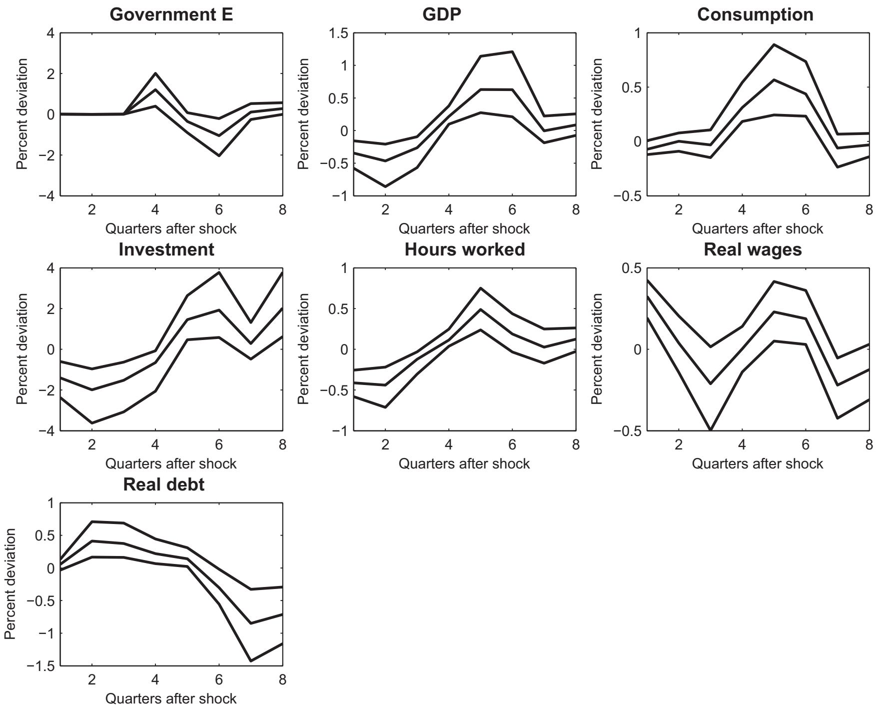


Figure B.5: Impulse Responses after a pre-announced increase in government expenditures. Pre-announcement three quarters. (68% probability bands)

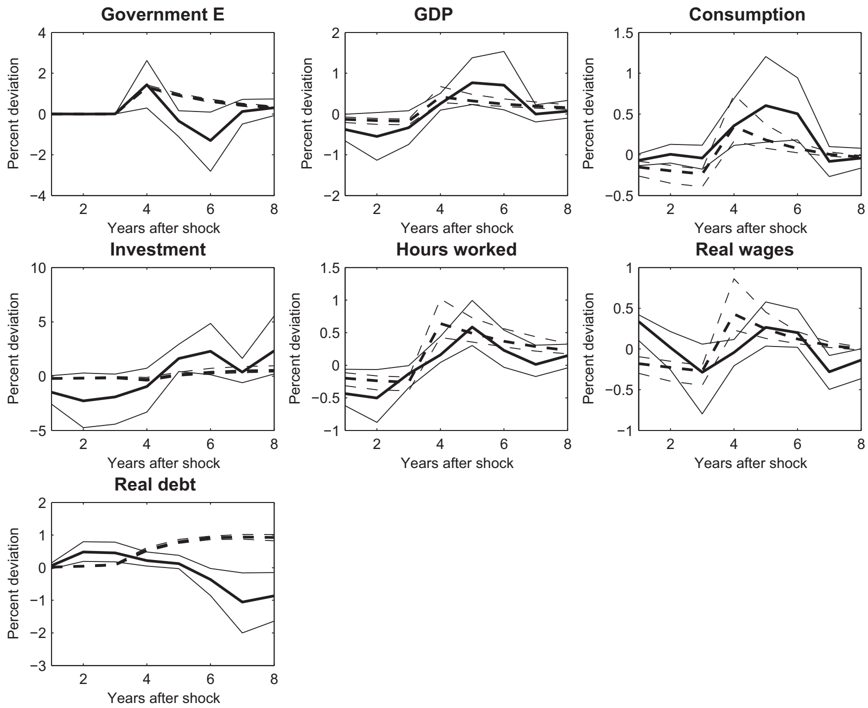


Figure B.6: Impulse Responses of the DSGE model (dashed) and the VMA model after a pre-announced (three quarters) increase in government expenditures. 68% probability bands

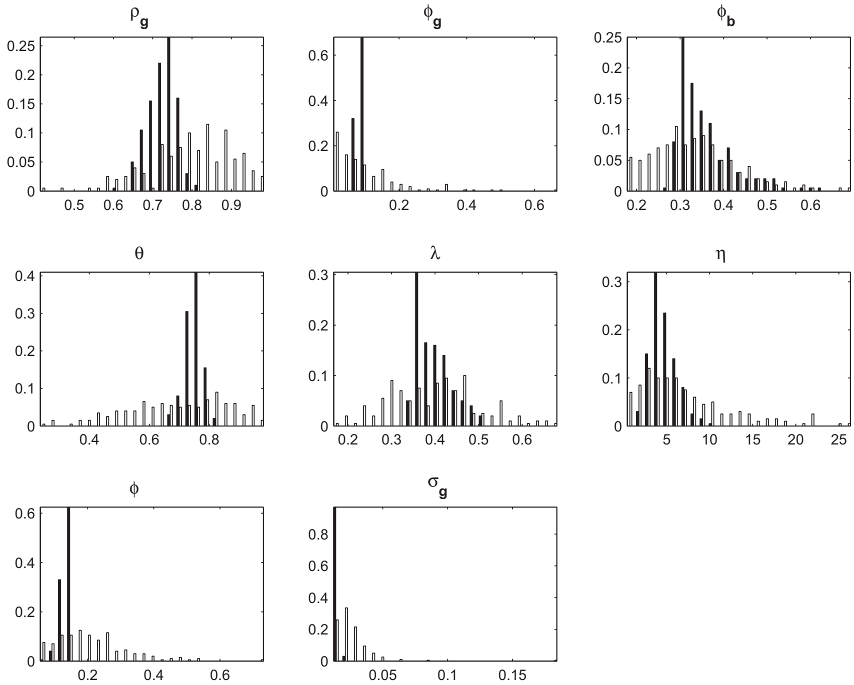


Figure B.7: Prior (white) vs. Posterior (black) distribution of the DSGE model

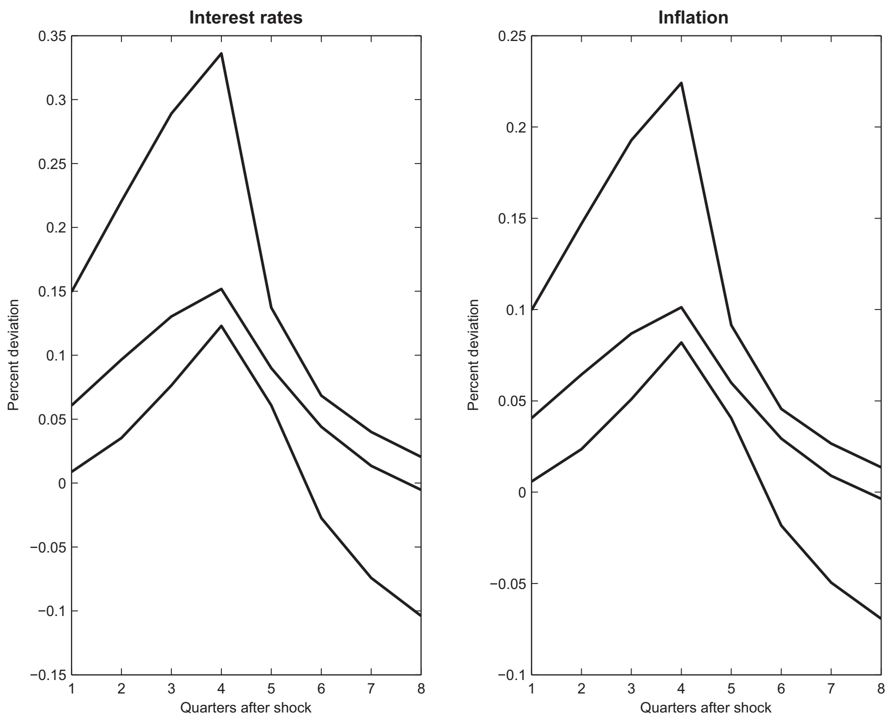


Figure B.8: Impulse responses of the nominal interest rate and inflation in the DSGE model.

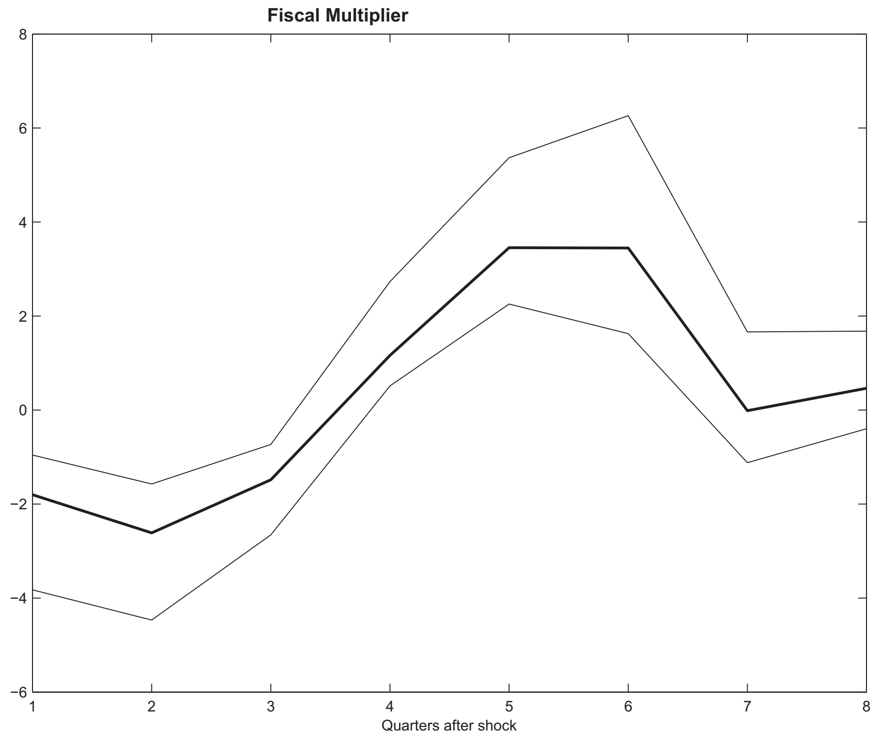


Figure B.9: Fiscal multiplier defined as fm=(response of gdp/change in government consumption expenditures)/average share of government consumption expenditures in gdp. 68% probability bands.

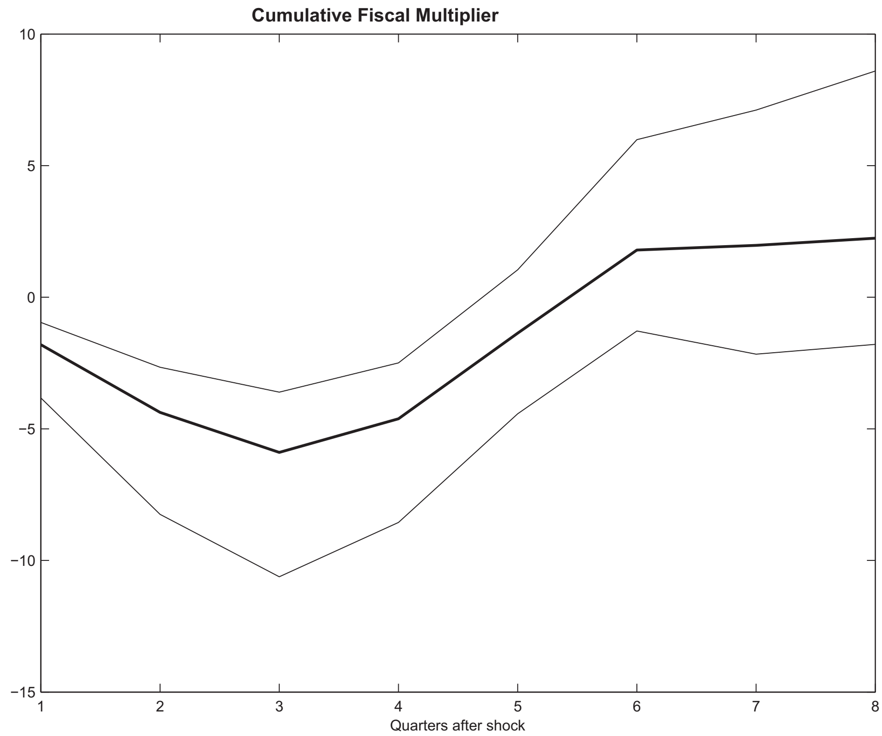


Figure B.10: Fiscal multiplier defined as fm=(cumulative response of gdp/change in government consumption expenditures)/average share of government consumption expenditures in gdp. 68% probability bands.

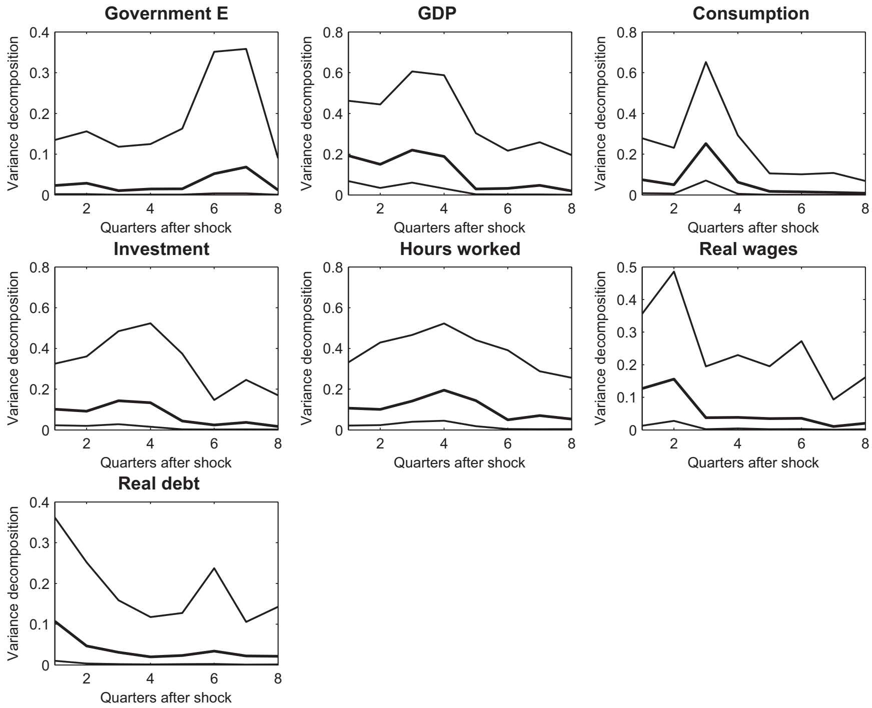


Figure B.11: Variance decomposition: Business cycle shock. 68% probability bands.

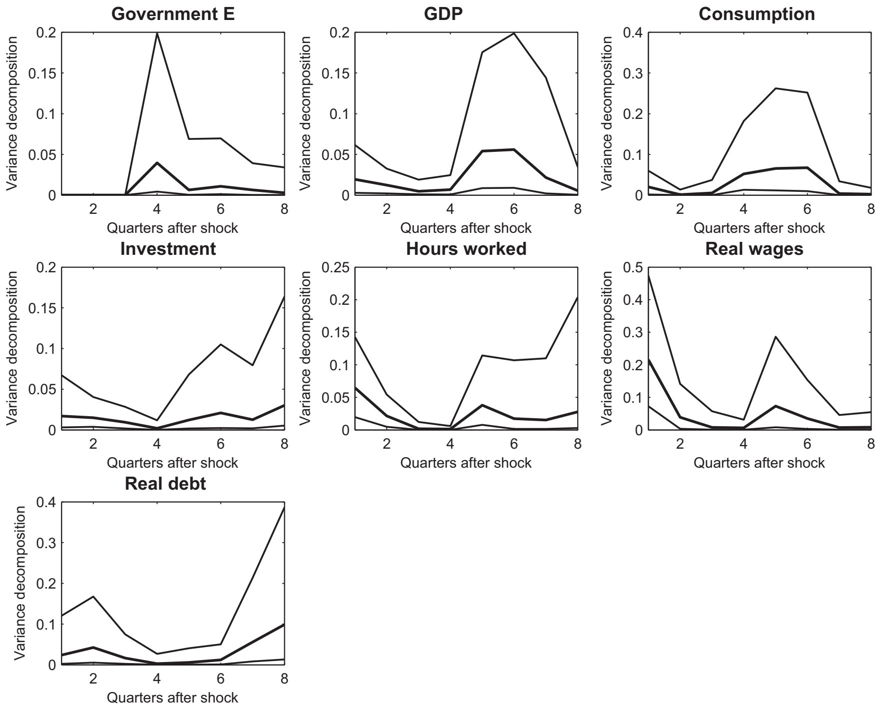


Figure B.12: Variance decomposition: Government expenditure shock. 68% probability bands.

Appendix C

Technical Appendix to chapter 4

C.1 Proof of proposition 1

The period utility function of the average household in equilibrium is given by:

$$\int_0^1 [u(\bullet) - v(l_{jt}) + z(m_t)]dj = u(y_t - g_t - \eta(y_{t-1} - g_{t-1})) - \int_0^1 v(l_{jt})dj + z(m_t).$$

To derive (4.27) we need to impose that, in the optimal steady state, real money balances are sufficiently close to being satiated (see Woodford, 2003a, Assumption 6.1) such that we can treat $(R-1)/R$ as an expansion parameter.

The first summand can be approximated to second order by:

$$\begin{aligned} u(y_t - g_t - \eta(y_{t-1} - g_{t-1})) = u_c y(1-\beta\eta)[\widehat{y}_t + \frac{1}{2}(1 - \varphi(1+\eta^2\beta))\widehat{y}_t^2 + \varphi\eta\widehat{y}_t\widehat{y}_{t-1} \\ + \varphi\widehat{y}_t(-\eta g_{t-1} - \beta\eta g_{t+1} + (1+\eta^2\beta)g_t)] + t.i.s.p. + \mathcal{O}(\|\widehat{\xi}_t\|^3), \quad \text{(C.1)} \end{aligned}$$

where we used $(x_t - x) = x(\widehat{x}_t + 0.5\widehat{x}_t^2) + \mathcal{O}(\|\widehat{x}_t\|^3)$, $\varphi = \frac{\sigma}{1-\beta\eta}$, t.i.s.p denotes terms independent of stabilization policy, $\sigma = -yu_{11}/u_1$, and $g_t = (G_t - G)/y$.

Since $y_t = a_t l_t/\Delta_t$, the second term can be approximated by

$$v(l_t) = u_c(1-\beta\eta)[\widehat{y}_t + \frac{1+\omega}{2}\widehat{y}_t^2 - (1+\omega)\widehat{a}_t\widehat{y}_t + \widehat{\Delta}_t] + t.i.s.p. + \mathcal{O}(\|\widehat{\xi}_t\|^3), \quad \text{(C.2)}$$

where we posited that in the equilibrium under flexible wages each household supplies the same amount of labor, $l = y$, $\omega = \frac{v_{ll}}{v_l}l$, and that due to the existence of an output subsidy the steady state is rendered efficient with $v_l = u_c(1-\beta\eta)$. In the next step we combine (C.1) and (C.2), employ $\tilde{g}_t = -\eta g_{t-1} - \beta\eta E_t g_{t+1} + (1+\eta^2\beta)g_t$, and obtain:

$$u(y_t - g_t - \eta(y_{t-1} - g_{t-1})) - \int_0^1 v(l_{jt})dj = u_c y(1-\beta\eta)[\frac{1}{2}(-\varphi(1+\eta^2\beta) - \omega)\widehat{y}_t^2 + \varphi\eta\widehat{y}_t\widehat{y}_{t-1} + \varphi\widehat{y}_t\tilde{g}_t + (1+\omega)\widehat{a}_t\widehat{y}_t - \widehat{\Delta}_t] + t.i.s.p. + \mathcal{O}(\|\widehat{\xi}_t\|^3). \quad \text{(C.3)}$$

The efficient rate of output is defined by the following difference equation:

$$[\omega + \varphi(1+\beta\eta^2)]\widehat{y}_t^e - \varphi\eta\widehat{y}_{t-1}^e - \varphi\eta\beta E_t\widehat{y}_{t+1}^e = \varphi\tilde{g}_t + (1+\omega)\widehat{a}_t + \mathcal{O}(\|\widehat{\xi}_t\|^2).$$

If we use this expression to rewrite (C.3), we obtain the following:

$$E_{t_0}\sum_{t=t_0}^{\infty}\beta^{t-t_0}\{u(\bullet) - \int_0^1 v(l_{jt})dj\} = -E_{t_0}\sum_{t=t_0}^{\infty}\beta^{t-t_0}u_c y(1-\beta\eta)\{\frac{1}{2}(\varphi(1+\eta^2\beta)+\omega)\widehat{y}_t^2 - \varphi\eta\widehat{y}_t\widehat{y}_{t-1} - [\omega+\varphi(1+\beta\eta^2)]\widehat{y}_t^e\widehat{y}_t + \varphi\eta\widehat{y}_{t-1}^e\widehat{y}_t + \varphi\eta\beta E_t\widehat{y}_{t+1}^e\widehat{y}_t + \widehat{\Delta}_t\} + t.i.s.p. + \mathcal{O}(\|\widehat{\xi}_t\|^3). \quad \text{(C.4)}$$

We seek to rewrite this expression in purely quadratic terms of the welfare-relevant gaps for inflation and output. To do so we apply the method of undetermined coefficients to reformulate the first part (all but $\widehat{\Delta}_t$), i.e. we seek to find the coefficient δ_0, such that (C.4) and

$$\begin{aligned}
&-\frac{1}{2}\delta_0(\widehat{y}_t - \widehat{y}_t^e - \delta^*(\widehat{y}_{t-1} - \widehat{y}_{t-1}^e))^2 \\
&= -\frac{1}{2}\delta_0 \\
&[\widehat{y}_t^2 - 2\widehat{y}_t\widehat{y}_t^e + (\widehat{y}_t^e)^2 - 2\delta^*(\widehat{y}_t - \widehat{y}_t^e)(\widehat{y}_{t-1} - \widehat{y}_{t-1}^e) + (\delta^*)^2(\widehat{y}_{t-1}^2 - 2\widehat{y}_{t-1}\widehat{y}_{t-1}^e + (\widehat{y}_{t-1}^e)^2)] \\
&= -\frac{1}{2}\delta_0[\widehat{y}_t^2 - 2\widehat{y}_t\widehat{y}_t^e - 2\delta^*\widehat{y}_t\widehat{y}_{t-1} + 2\delta^*\widehat{y}_t\widehat{y}_{t-1}^e + 2\delta^*\widehat{y}_t^e\widehat{y}_{t-1} + (\delta^*)^2\widehat{y}_{t-1}^2 - 2(\delta^*)^2\widehat{y}_{t-1}\widehat{y}_{t-1}^e] \\
&= -\frac{1}{2}\delta_0[((\delta^*)^2\beta + 1)\widehat{y}_t^2 - 2\delta^*\widehat{y}_t\widehat{y}_{t-1} + 2\delta^*\widehat{y}_t\widehat{y}_{t-1}^e + 2\delta^*\beta\widehat{y}_{t+1}^e\widehat{y}_t - (2(\delta^*)^2\beta + 2)\widehat{y}_t\widehat{y}_t^e] \\
&= -\frac{1}{2}\delta_0((\delta^*)^2\beta + 1)\widehat{y}_t^2 + \delta_0\delta^*\widehat{y}_t\widehat{y}_{t-1} - \delta_0\delta^*\widehat{y}_t\widehat{y}_{t-1}^e - \delta_0\delta^*\beta\widehat{y}_{t+1}^e\widehat{y}_t + \delta_0((\delta^*)^2\beta + 1)\widehat{y}_t\widehat{y}_t^e
\end{aligned}$$

are consistent. We use that $\widehat{y}_{t0-1}$ is *t.i.s.p.*. The parameter δ^*, $0 \leq \delta^* \leq \eta$, is the smaller root of this quadratic equation: $\eta\varphi(1+\beta\delta^2) = [\omega + \varphi(1+\beta\eta^2)]\delta$. This root is assigned to past values of the natural and efficient rate of output in their stationary solutions. Comparing coefficients, δ_0 is

$$\delta_0 = \frac{u_c y(1-\beta\eta)\eta\varphi}{\delta^*}.$$

If firms are allowed to index with past inflation, such that

$$E_{t_0}\sum_{t=t_0}^{\infty}\beta^{t-t_0}2\widehat{\Delta}_t = E_{t_0}\sum_{t=t_0}^{\infty}\beta^{t-t_0}\frac{\zeta\alpha}{(1-\alpha)(1-\alpha\beta)}(\widehat{\pi}_t-\gamma\widehat{\pi}_{t-1})^2 + t.i.s.p. + \mathcal{O}(\|\widehat{\xi}_t\|^3),$$

the quadratic approximation in (C.4) can be written as:

$$\begin{aligned}&-E_{t_0}\sum_{t=t_0}^{\infty}\beta^{t-t_0}\frac{u_c y(1-\beta\eta)}{2}\\ &[\frac{\eta\varphi}{\delta^*}(\widehat{y}_t-\widehat{y}_t^e-\delta^*(\widehat{y}_{t-1}-\widehat{y}_{t-1}^e))^2+\frac{\zeta\alpha}{(1-\alpha)(1-\alpha\beta)}(\widehat{\pi}_t-\gamma\widehat{\pi}_{t-1})^2] + t.i.s.p. + \mathcal{O}(\|\widehat{\xi}_t\|^3).\end{aligned}$$

The last approximation needed is that involving the utility of real money balances. Applying similar techniques we get

$$z(m_t) = z + yu_c(s_m(\widehat{m}_t + \frac{1}{2}s_m(1-\sigma_m)\widehat{m}_t^2) + t.i.s.p + \mathcal{O}(\|\widehat{\xi}_t\|^3), \tag{C.5}$$

where we employ $s_m = z_m m/(u_c y) = (R-1)(1-\beta\eta)R$ and $\sigma_m = -z_{mm}m/z_m$. Since we treat $(R-1)/R$ as an expansion parameter, s_m and $1/\sigma_m$ are of first order. However, $s_m\sigma_m$ approaches a finite limit for $(R-1)/R \to 0$, which is given by

$$s_m\sigma_m = \frac{z_{mm}m^2}{yu_c}.$$

The interest elasticity of money demand is given by the following expression:

$$\eta_i = -\frac{u_c(1-\beta_n)}{z_{mm}}\frac{1-\frac{R-1}{R}}{m} = \frac{1}{\sigma_m(R-1)}.$$

At the limit for $(R-1)/R \to 0$, it follows that $\eta_i = -u_c(1-\beta\eta)/(z_{mm}m)$ and therefore $s_m\sigma_m = (1-\beta\eta)/(v\eta_i)$, with $v = y/m$. A first-order approximation of the money demand equation (4.24) yields

$$\widehat{m}_t = -\eta_i\widehat{R}_t - \frac{1}{\sigma_m}\widehat{\lambda}_t + \mathcal{O}(\|\widehat{\xi}_t|^2),$$

where

$$\widehat{\lambda}_t = -\varphi(\widehat{y}_t - \eta\widehat{y}_{t-1}) + \beta\eta\varphi(\widehat{y}_{t+1} - \eta\widehat{y}_t) + \varphi(g_t - \eta g_{t-1}) - \beta\eta\varphi(g_{t+1} - \eta g_t) + \mathcal{O}(\|\widehat{\xi}_t|^2).$$

Using all the above we can rewrite $z(m_t)$ in the following way:

$$z(m_t) = -\frac{\eta_i y u_c}{2v}(1-\beta\eta)(\widehat{R}_t^2 + 2\frac{R-1}{R}\widehat{R}_t) + t.i.s.p + \mathcal{O}(\|\widehat{\xi}_t, (R-1)/R\|^3). \quad \text{(C.6)}$$

We assume for simplicity that $[(R-1)/R - 0]$ is of second order, and sum the results in expression (4.27) in the text.

C.2 Estimation Results

Table C.1: Prior distribution of the structural parameters

	Prior distribution		
Parameter	distribution	mean	std
ρ	beta	0.8	0.1
ϕ_π	normal	1.7	0.1
ϕ_y	normal	0.125	0.05
ω	gamma	1	0.5
σ_c	normal	1.5	0.375
α	beta	0.75	0.05
η	beta	0.7	0.1
γ	beta	0.75	0.15
σ_m	normal	1.25	0.375
ψ_g	beta	0.7	0.1
ψ_a	beta	0.7	0.1
σ_g	invgamma	0.04	0.026
σ_a	invgamma	0.04	0.026
σ_μ	invgamma	0.04	0.026

Figure C.1: Deep parameters prior vs. posterior (black) distribution in Model 1

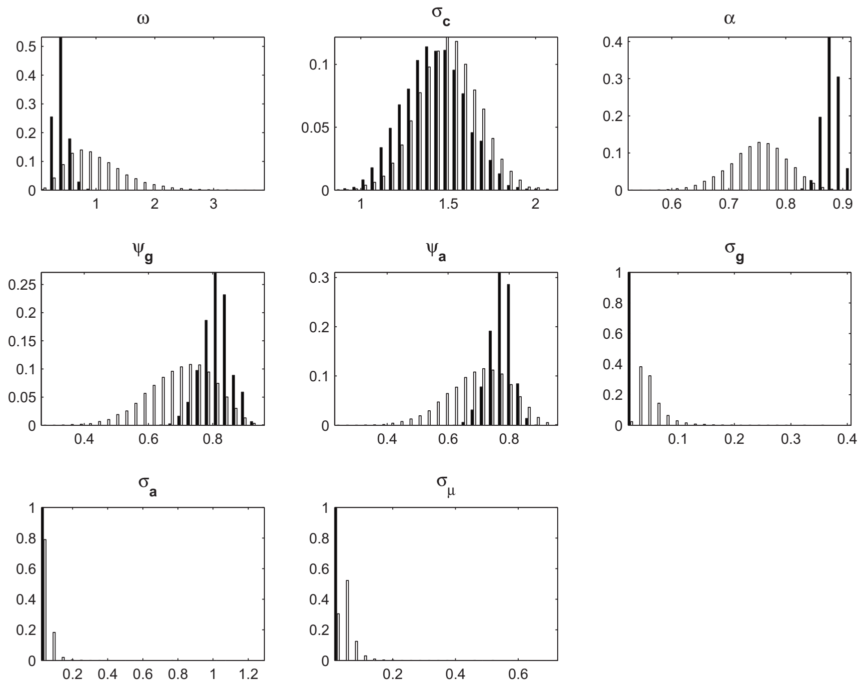


Figure C.2: Deep parameters prior vs. posterior (black) distribution in Model 2

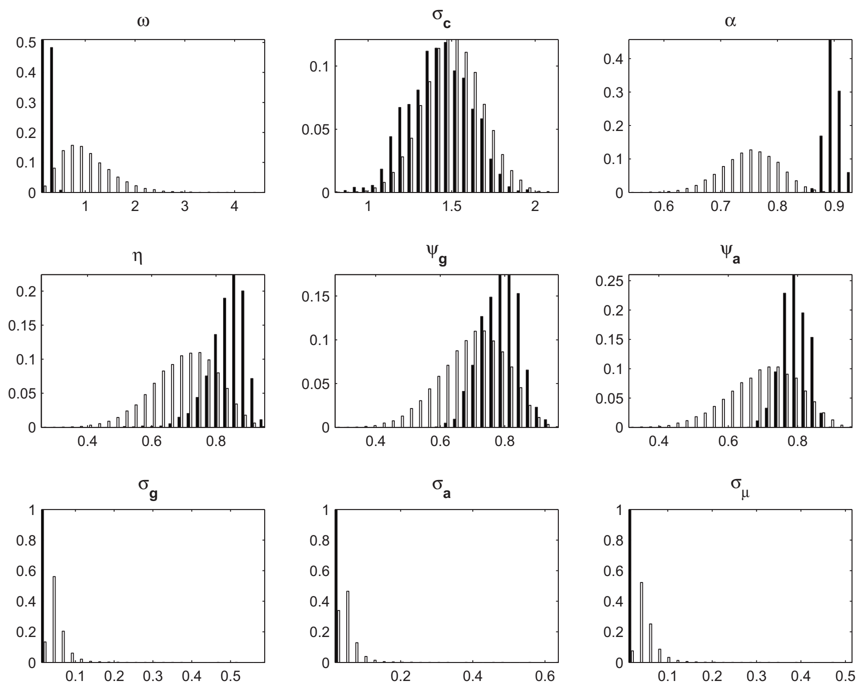


Table C.2: Posterior estimates of the structural parameters in each model

	$Model_1$		$Model_2$		$Model_3$		$Model_4$		$Model_5$	
Parameter	mean	std	mean	std	mean	std	mean	std	mean	std
ρ	0.4379	0.0719	0.4582	0.0727	0.3578	0.0742	0.4477	0.0730	0.3756	0.0756
ϕ_π	1.6972	0.0983	1.7188	0.0985	1.6255	0.1009	1.6719	0.0963	1.6772	0.0972
ϕ_y	0.0964	0.0263	0.0701	0.0261	0.1154	0.0324	0.0939	0.0261	0.0821	0.0297
ω	0.4170	0.1104	0.2610	0.0547	0.4156	0.0984	0.3962	0.0903	0.2904	0.0644
σ_c	1.4202	0.1744	1.4296	0.1798	1.4142	0.1879	1.4117	0.1735	1.4158	0.1779
α	0.8810	0.0139	0.8988	0.0126	0.8672	0.0177	0.8855	0.0156	0.8777	0.0168
η	-	-	0.8368	0.0566	-	-	-	-	0.8307	0.0515
γ	-	-	-	-	0.4677	0.0842	-	-	0.4563	0.0796
σ_m	-	-	-	-	-	-	1.2685	0.3866	1.2230	0.3777
ψ_g	0.8126	0.0438	0.7880	0.0579	0.7992	0.0482	0.8099	0.0477	0.7950	0.0571
ψ_a	0.7733	0.0367	0.7929	0.0372	0.6766	0.0545	0.7836	0.0403	0.6780	0.0562
σ_g	0.0095	0.0007	0.0084	0.0005	0.0097	0.0007	0.0095	0.0007	0.0084	0.0005
σ_a	0.0168	0.0030	0.0206	0.0034	0.0178	0.0032	0.0174	0.0026	0.0201	0.0039
σ_μ	0.0092	0.0006	0.0090	0.0006	0.0097	0.0007	0.0092	0.0006	0.0094	0.0006

Figure C.3: Deep parameters prior vs.posterior (black) distribution in Model 3

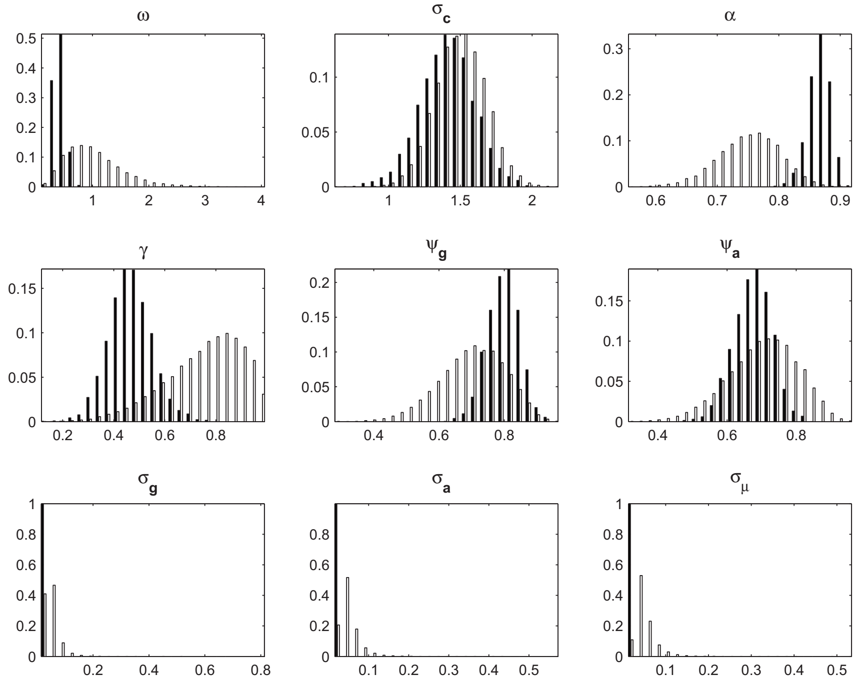


Figure C.4: Deep parameters prior vs. posterior (black) distribution in Model 4

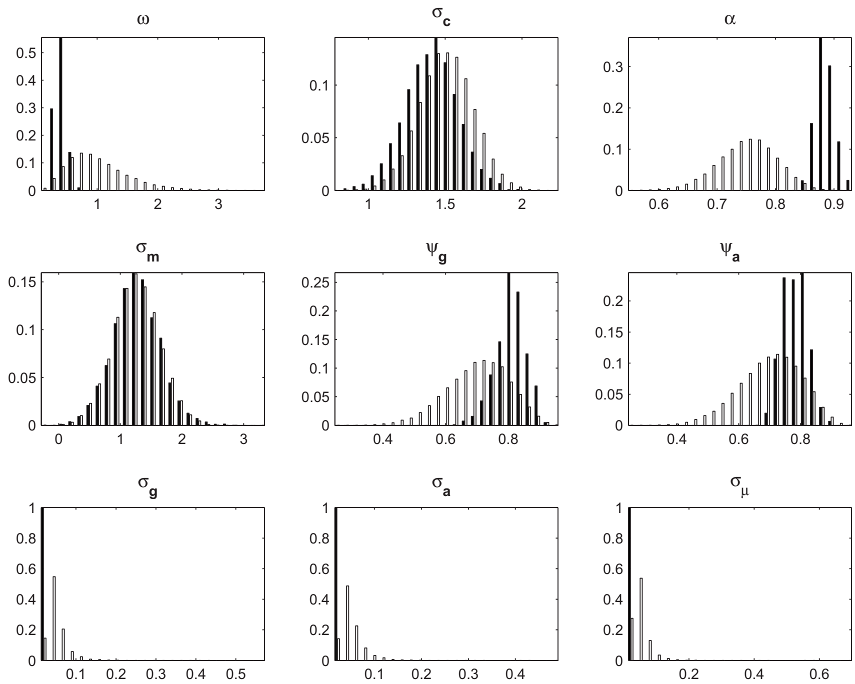
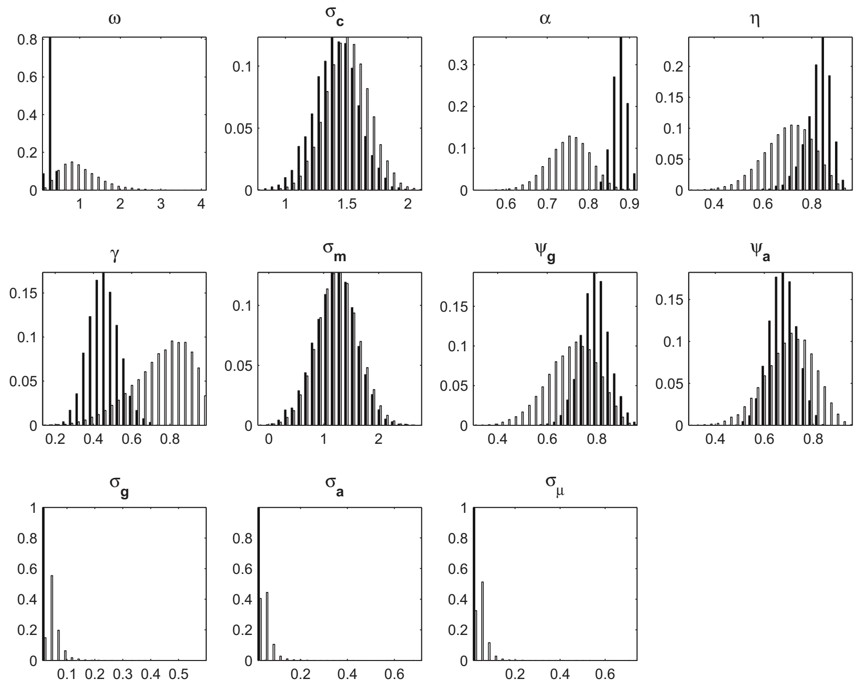


Figure C.5: Deep parameters prior vs. posterior (black) distribution in Model 5



Bibliography

M. Adolfson, S. Laseen, J. Linde, and L. E. Svensson. Optimal monetary policy in an operational medium-sized dsge model. NBER Working Papers 14092, National Bureau of Economic Research, Inc, June 2008.

S. R. Aiyagari, L. J. Christiano, and M. Eichenbaum. The output, employment, and interest rate effects of government consumption. Discussion Paper / Institute for Empirical Macroeconomics 25, Federal Reserve Bank of Minneapolis, Apr. 1990.

D. Altig, L. J. Christiano, M. Eichenbaum, and J. Linde. Technology shocks and aggregate fluctuations. draft, Northwestern University, 2002.

S. Altug. Time-to-build and aggregate fluctuations: Some new evidence. *International Economic Review*, 30(4):889–920, November 1989.

S. An and F. Schorfheide. Bayesian analysis of dsge models by s. an and f. schorfheide. *Econometric Reviews*, 26(2):187–192, 2007.

R. J. Barro. Output effects of government purchases. *Journal of Political Economy*, 89(6):1086–1121, December 1981.

R. J. Barro. Government spending, interest rates, prices, and budget deficits in the united kingdom, 1701-1918. *Journal of Monetary Economics*, 20(2):221–247, September 1987.

M. Baxter and R. G. King. Fiscal policy in general equilibrium. *American Economic Review*, 83(3):315–34, June 1993.

A. Beyer and R. E. Farmer. Identification problems in sdge models with an illustration to a small macro model. Computing in Economics and Finance 2006 81, Society for Computational Economics, July 2006.

O. Blanchard and R. Perotti. An empirical characterization of the dynamic effects of changes in government spending and taxes on output. *The Quarterly Journal of Economics*, 117(4):1329–1368, November 2002.

O. J. Blanchard and D. Quah. The dynamic effects of aggregate demand and supply disturbances. *American Economic Review*, 79(4):655–73, September 1989.

H. Bohn. Budget balance through revenue or spending adjustments? : Some historical evidence for the united states. *Journal of Monetary Economics*, 27(3):333–359, June 1991.

J. Boivin and M. Giannoni. Has monetary policy become more effective? *Review of Economics and Statistics*, 3(88):445–462, 2006.

W. Brainard. Uncertainty and the effectiveness of policy. *American Economic Review*, 57:411–425, 1967.

W. Brock, S. Durlauf, and K. West. Model uncertainty and policy evaluation: some theory and empirics. *Proceedings*, 2005.

C. Burnside, M. Eichenbaum, and J. D. M. Fisher. Fiscal shocks and their consequences. *Journal of Economic Theory*, 115(1):89–117, March 2004.

G. A. Calvo. Staggered price setting in a utility-maximizing framework. *Journal of Monetary Economics*, 12:383–398, 1983.

F. Canova. Statistical inference in calibrated models. *Journal of Applied Econometrics*, 9(S):S123–44, Suppl. De 1994.

F. Canova. Sensitivity analysis and model evaluation in simulated dynamic general equilibrium economies. *International Economic Review*, 36(2):477–501, May 1995.

F. Canova and G. D. Nicolo. Monetary disturbances matter for business fluctuations in the g-7. *Journal of Monetary Economics*, 49(6):1131–1159, September 2002.

F. Canova and J. P. Pina. Monetary policy misspecification in var models. CEPR Discussion Papers 2333, C.E.P.R. Discussion Papers, Dec. 1999.

F. Canova and L. Sala. Back to square one: Identification issues in dsge models. Working Papers 303, IGIER (Innocenzo Gasparini Institute for Economic Research), Bocconi University, 2006.

L. J. Christiano and M. Eichenbaum. Current real-business-cycle theories and aggregate labor-market fluctuations. *American Economic Review*, 82(3):430–50, June 1992.

L. J. Christiano, M. Eichenbaum, and C. L. Evans. Monetary policy shocks: What have we learned and to what end? In J. B. Taylor and M. Woodford, editors, *Handbook of Macroeconomics*, volume 1 of *Handbook of Macroeconomics*, chapter 2, pages 65–148. Elsevier, January 1999.

L. J. Christiano, M. Eichenbaum, and C. L. Evans. Nominal rigidities and the dynamic effects of a shock to monetary policy. *Journal of Political Economy*, 113(1):1–45, February 2005.

L. J. Christiano, M. Trabandt, and K. Walentin. Introducing financial frictions and unemployment into a small open economy model. Working Paper Series 214, Sveriges Riksbank (Central Bank of Sweden), Nov. 2007.

H. Chung and E. M. Leeper. What has financed government debt? NBER Working Papers 13425, National Bureau of Economic Research, Inc, Sept. 2007.

G. Coenen. Inflation persistence and robust monetary policy design. *Journal of Economic Dynamics and Control*, 31(1):111–140, January 2007.

G. Coenen, A. Levin, and V. Wieland. Data uncertainty and the role of money as an information variable for monetary policy. *European Economic Review*, 49(4): 975–1006, May 2005.

T. Cogley and J. M. Nason. Output dynamics in real-business-cycle models. *American Economic Review*, 85(3):492–511, June 1995.

T. Cogley, R. Colacito, and T. J. Sargent. Benefits from u.s. monetary policy experimentation in the days of samuelson and solow and lucas. *Journal of Money, Credit and Banking*, 39(s1):67–99, 02 2007.

T. Cogley, R. Colacito, L. P. Hansen, and T. J. Sargent. Robustness and u.s. monetary policy experimentation. mimeo, New York University, 2008.

G. Corsetti, A. Meier, and G. J. Müller. Fiscal stimulus with spending reversals. CEPR Discussion Papers 7302, C.E.P.R. Discussion Papers, May 2009. URL `http://ideas.repec.org/p/cpr/ceprdp/7302.html`.

T. Davig and E. M. Leeper. Fluctuating macro policies and the fiscal theory. NBER Working Papers 11212, National Bureau of Economic Research, Inc, Mar. 2005.

D. N. DeJong, B. F. Ingram, and C. H. Whiteman. A bayesian approach to dynamic macroeconomics. *Journal of Econometrics*, 98(2):203–223, October 2000.

M. Del Negro and F. Schorfheide. Forming priors for dsge models (and how it affects the assessment of nominal rigidities). NBER Working Papers 13741, National Bureau of Economic Research, 2008.

M. DelNegro and F. Schorfheide. Priors from general equilibrium models for vars. *International Economic Review*, 45(2):643–673, 05 2004.

M. B. Devereux, A. C. Head, and B. J. Lapham. Monopolistic competition, increasing returns, and the effects of government spending. *Journal of Money, Credit and Banking*, 28(2):233–54, May 1996.

T. Doan, R. Litterman, and C. Sims. Forecasting and conditional projection using realistic prior distributions. *Econometric Reviews*, 3(1):1–100, 1984.

M. Dwyer. Impulse response priors for discriminating structural vector autoregressions. UCLA Economics Working Papers 780, UCLA Department of Economics, Aug. 1998.

W. Edelberg, M. Eichenbaum, and J. D. Fisher. Understanding the effects of a shock to government purchases. *Review of Economic Dynamics*, 2(1):166–206, January 1999.

G. Fagan, J. Henry, and R. Mestre. An area-wide model (awm) for the euro area. Technical Report 42, ECB, 2001.

J. Faust. The robustness of identified var conclusions about money. International Finance Discussion Papers 610, Board of Governors of the Federal Reserve System (U.S.), 1998.

L. Forni, L. Monteforte, and L. Sessa. The general equilibrium effects of fiscal policy: Estimates for the euro area. *Journal of Public Economics*, 93(3-4):559–585, April 2009.

N. Francis and V. A. Ramey. Is the technology-driven real business cycle hypothesis dead? shocks and aggregate fluctuations revisited. *Journal of Monetary Economics*, 52(8):1379–1399, November 2005.

F. Furlanetto and M. Seneca. Fiscal shocks and real rigidities. *Topics in Macroeconomics*, 9(1):1781–1781, 2009.

J. Galí. Technology, employment, and the business cycle: Do technology shocks explain aggregate fluctuations? *American Economic Review*, 89(1):249–271, March 1999.

J. Galí, J. López-Salido, and J. Vallés. Understanding the effects of government spending on consumption. *Journal of the European Economic Association*, 5(1):227–270, 03 2007.

J. Geweke. Computational experiments and reality. Technical report, University of Minnesota and Federal Reserve Bank of Minneapolis, June 1999a.

J. Geweke. Using simulation methods for bayesian econometric models. Computing in economics and finance 1999, Society for Computational Economics, Mar. 1999b.

M. P. Giannoni. Robust optimal monetary policy in a forward-looking model with parameter and shock uncertainty. *Journal of Applied Econometrics*, 22(1):179–213, 2007.

G. J. Hall. Overtime, effort, and the propagation of business cycle shocks. *Journal of Monetary Economics*, 38(1):139–160, August 1996.

R. E. Hall. Labor supply and aggregate fluctuations. NBER Working Papers 0385, National Bureau of Economic Research, Inc, June 1980.

L. P. Hansen. Large sample properties of generalized method of moments estimators. *Econometrica*, 50(4):1029–54, July 1982.

L. P. Hansen and J. J. Heckman. The empirical foundations of calibration. *Journal of Economic Perspectives*, 10(1):87–104, Winter 1996.

L. P. Hansen and T. J. Sargent. Two difficulties in interpreting vector autoregressions. In L. P. Hansen and T. J. Sargent, editors, *Rational Expectations Econometrics*. Westview Press, 1991.

L. P. Hansen and T. J. Sargent. Robust control and model uncertainty. *American Economic Review*, 91(2):60–66, May 2001a.

L. P. Hansen and T. J. Sargent. Acknowledging misspecification in macroeconomic theory. *Review of Economic Dynamics*, 4(3):519–535, July 2001b.

L. P. Hansen and T. J. Sargent. Robust control of forward-looking models. *Journal of Monetary Economics*, 50(3):581–604, April 2003.

L. P. Hansen, T. J. Sargent, and J. Tallarini, Thomas D. Robust permanent income and pricing. *Review of Economic Studies*, 66(4):873–907, October 1999.

R. J. Hodrick and E. C. Prescott. Postwar u.s. business cycles: An empirical investigation. *Journal of Money, Credit and Banking*, 29(1):1–16, February 1997.

C. L. House and M. D. Shapiro. Phased-in tax cuts and economic activity. *American Economic Review*, 96(5):1835–1849, December 2006.

B. F. Ingram and C. H. Whiteman. Supplanting the 'minnesota' prior: Forecasting macroeconomic time series using real business cycle model priors. *Journal of Monetary Economics*, 34(3):497–510, December 1994.

P. N. Ireland. Sticky-price models of the business cycle: Specification and stability. *Journal of Monetary Economics*, 47(1):3–18, February 2001.

P. N. Ireland. A method for taking models to the data. *Journal of Economic Dynamics and Control*, 28(6):1205–1226, March 2004.

R. J. Tetlow and P. von zur Muehlen. Robust monetary policy with misspecified models: Does model uncertainty always call for attenuated policy? *Journal of Economic Dynamics and Control*, 25(6-7):911–949, June 2001.

J. Kim. Constructing and estimating a realistic optimizing model of monetary policy. *Journal of Monetary Economics*, 45(2):329–359, April 2000.

A. Kociecki. Priors for impulse responses in bayesian structural var models. Technical report, National Bank of Poland, 2005.

K. Küster and V. Wieland. Insurance policies for monetary policy in the euro area, Mar. 2005.

F. E. Kydland. On the econometrics of world business cycles. *European Economic Review*, 36(2-3):476–482, April 1992.

F. E. Kydland and E. C. Prescott. Time to build and aggregate fluctuations. *Econometrica*, 50(6):1345–70, November 1982.

F. E. Kydland and E. C. Prescott. The econometrics of the general equilibrium approach to business cycles. *Scandinavian Journal of Economics*, 93(2):161–78, 1991.

M. Lanne and H. Lütkepohl. Structural vector autoregressions with nonnormal residuals. Economics Working Papers ECO2005/25, European University Institute, 2005.

M. Lanne and H. Lütkepohl. Identifying monetary policy shocks via changes in volatility. *Journal of Money, Credit and Banking*, 40(6):1131–1149, 09 2008.

M. Lanne, H. Lütkepohl, and K. Maciejowska. Structural vector autoregressions with markov switching. Economics Working Papers ECO2009/06, European University Institute, 2009.

E. Leamer. *Specification Searches.* John Wiley and Sons, 1978.

B.-S. Lee and B. F. Ingram. Simulation estimation of time-series models. *Journal of Econometrics*, 47(2-3):197–205, February 1991.

E. M. Leeper. Equilibria under 'active' and 'passive' monetary and fiscal policies. *Journal of Monetary Economics*, 27(1):129–147, February 1991.

E. M. Leeper and S.-C. S. Yang. Dynamic scoring: Alternative financing schemes. *Journal of Public Economics*, 92(1-2):159–182, February 2008.

E. M. Leeper, T. B. Walker, and S.-C. S. Yang. Fiscal foresight: Analytics and econometrics. NBER Working Papers 14028, National Bureau of Economic Research, Inc, May 2008.

A. Levin, V. Wieland, and J. C. Williams. The performance of forecast-based monetary policy rules under model uncertainty. *American Economic Review*, 93(3):622–645, June 2003.

A. Levin, A. Onatski, J. Williams, and N. Williams. Monetary policy under uncertainty in microfounded macroeconometric models. *NBER Macroeconomics Annual*, 2005.

A. T. Levin and J. C. Williams. Robust monetary policy with competing reference models. *Journal of Monetary Economics*, 50(5):945–975, July 2003.

P. Levine, P. McAdam, J. Pearlman, and R. Pierse. Risk management in action. robust monetary policy rules under structured uncertainty. Working Paper Series 870, European Central Bank, Feb. 2008.

M. Lippi and L. Reichlin. The dynamic effects of aggregate demand and supply disturbances: Comment. *American Economic Review*, 83(3):644–52, June 1993.

M. Lippi and L. Reichlin. Var analysis, nonfundamental representations, blaschke matrices. *Journal of Econometrics*, 63(1):307–325, July 1994.

T. A. Lubik and F. Schorfheide. Testing for indeterminacy: An application to u.s. monetary policy. *American Economic Review*, 94(1):190–217, March 2004.

R. E. Lucas. Macroeconomic priorities. *American Economic Review*, 93(1):1–14, March 2003.

B. T. Mccallum. Robustness properties of a rule for monetary policy. *Carnegie-Rochester Conference Series on Public Policy*, 29:173–203, 1988.

E. McGrattan, V. V. Chari, and P. Kehoe. Are structural vars useful guides for developing business cycle theories? 2005 Meeting Papers 664, Society for Economic Dynamics, 2005.

E. R. McGrattan. The macroeconomic effects of distortionary taxation. *Journal of Monetary Economics*, 33(3):573–601, June 1994.

E. R. McGrattan, R. Rogerson, and R. Wright. An equilibrium model of the business cycle with household production and fiscal policy. *International Economic Review*, 38(2):267–90, May 1997.

K. Mertens and M. O. Ravn. The aggregate effects of anticipated and unanticipated u.s. tax policy shocks: Theory and empirical evidence. Economics Working Papers ECO2008/05, European University Institute, 2008.

K. Mertens and M. O. Ravn. Measuring the impact of fiscal policy in the face of anticipation: A structural var approach. manuscript, European University Institute, 2009.

A. Mountford and H. Uhlig. What are the effects of fiscal policy shocks? SFB 649 Discussion Papers SFB649DP2005-039, Sonderforschungsbereich 649, Humboldt University, Berlin, Germany, July 2005.

A. Mountford and H. Uhlig. What are the effects of fiscal policy shocks? NBER Working Papers 14551, National Bureau of Economic Research, Inc, Dec. 2008.

M. D. Negro and F. Schorfheide. Monetary policy analysis with potentially misspecified models. *American Economic Review, forthcoming*, 2009.

A. Onatski and J. H. Stock. Robust monetary policy under model uncertainty in a small model of the us economy. *Macroeconomic Dynamics*, 6:85–110, 2002.

A. Onatski and N. Williams. Modeling model uncertainty. *Journal of the European Economic Association*, 1(5):1087–1122, 09 2003.

A. Orphanides and J. C. Williams. Robust monetary policy rules with unknown natural rates. *Brookings Papers on Economic Activity*, 33(2002-2):63–146, 2002.

A. Orphanides and J. C. Williams. Robust monetary policy with imperfect knowledge. *Journal of Monetary Economics*, 54(5):1406–1435, July 2007.

E. Pappa. New-keynesian or rbc transmission? the effects of fiscal shocks in labour markets. CEPR Discussion Papers 5313, C.E.P.R. Discussion Papers, Oct. 2005.

R. Perotti. In search of the transmission mechanism of fiscal policy. NBER Working Papers 13143, National Bureau of Economic Research, Inc, June 2007.

D. J. Poirier. Revising beliefs in nonidentified models. *Econometric Theory*, 14(04): 483–509, 1998.

V. A. Ramey. Identifying government spending shocks: It's all in the timing. Technical report, University of California, San Diego, 2008.

V. A. Ramey and M. D. Shapiro. Costly capital reallocation and the effects of government spending. *Carnegie-Rochester Conference Series on Public Policy*, 48:145–194, June 1998.

M. O. Ravn, S. Schmitt-Grohé, and M. Uribe. Explaining the effects of government spending shocks on consumption and the real exchange rate. NBER Working Papers 13328, National Bureau of Economic Research, Inc, Aug. 2007.

M. O. Ravn, S. Schmitt-Grohe, M. Uribe, and L. Uuskula. Deep habits and the dynamic effects of monetary policy shocks. Economics Working Papers ECO2008/40, European University Institute, 2008.

R. Rigobon. Identification through heteroskedasticity. *The Review of Economics and Statistics*, 85(4):777–792, 09 2003.

C. D. Romer and D. H. Romer. Does monetary policy matter? a new test in the spirit of friedman and schwartz. NBER Working Papers 2966, National Bureau of Economic Research, Inc, Feb. 1990.

J. Rotemberg and M. Woodford. An optimization-based econometric framework for the evaluation of monetary policy. In *NBER Macroeconomics Annual 1997, Volume 12*, NBER Chapters, pages 297–361. National Bureau of Economic Research, Inc, 1997.

J. Rotemberg and M. Woodford. Interest-rate-rules in an estimated sticky-price-model. In J. Taylor, editor, *Monetary Policy Rules*. University of Chicago Press, Chicago, 1999.

J. J. Rotemberg and M. Woodford. Oligopolistic pricing and the effects of aggregate demand on economic activity. *Journal of Political Economy*, 100(6):1153–1207, December 1992.

J. J. Rotemberg and M. Woodford. Interest-rate rules in an estimated sticky price model. NBER Working Papers 6618, National Bureau of Economic Research, Inc, June 1998.

J. F. Rubio-Ramirez, D. Waggoner, and T. Zha. Markov-switching structural vector autoregressions: theory and application. Working Paper 2005-27, Federal Reserve Bank of Atlanta, 2005.

G. D. Rudebusch. Assessing nominal income rules for monetary policy with model and data uncertainty. *Economic Journal*, 112(479):402–432, April 2002.

T. J. Sargent. Two models of measurements and the investment accelerator. *Journal of Political Economy*, 97(2):251–87, April 1989.

C. A. Sims. Macroeconomics and reality. *Econometrica*, 48(1):1–48, January 1980.

C. A. Sims. Pitfalls of a minimax approach to model uncertainty. *American Economic Review*, 91(2):51–54, May 2001.

C. A. Sims. Making macro models behave reasonably. Working Papers 74, Princeton University, Department of Economics, December 2006b.

C. A. Sims and T. Zha. Does monetary policy generate recessions? *Macroeconomic Dynamics*, 10(02):231–272, April 2006.

F. Smets and R. Wouters. An estimated dynamic stochastic general equilibrium model of the euro area. *Journal of the European Economic Association*, 1(5):1123–1175, 2003.

F. Smets and R. Wouters. Shocks and frictions in us business cycles: A bayesian dsge approach. *American Economic Review*, 97(3):586–606, 2007.

J. Smith, A A. Estimating nonlinear time-series models using simulated vector autoregressions. *Journal of Applied Econometrics*, 8(S):S63–84, Suppl. De 1993.

U. Soderstrom. Monetary policy with uncertain parameters. *Scandinavian Journal of Economics*, 104:125–145, 2002.

L. E. O. Svensson and M. Woodford. Indicator variables for optimal policy under asymmetric information. *Journal of Economic Dynamics and Control*, 28(4):661–690, January 2004.

J. Tenhofen and G. B. Wolff. Does anticipation of government spending matter? evidence from an expectation augmented var. Discussion Paper Series 1: Economic Studies 2007,14, Deutsche Bundesbank, Research Centre, 2007.

M. Trabandt and H. Uhlig. How far are we from the slippery slope? the laffer curve revisited. SFB 649 Discussion Papers SFB649DP2006-023, Sonderforschungsbereich 649, Humboldt University, Berlin, Germany, Apr. 2006.

S. J. Turnovsky and W. H. Fisher. The composition of government expenditure and its consequences for macroeconomic performance. *Journal of Economic Dynamics and Control*, 19(4):747–786, May 1995.

H. Uhlig. Bayesian vector autoregressions with stochastic volatility. *Econometrica*, 65 (1):59–74, January 1997.

H. Uhlig. What are the effects of monetary policy on output? results from an agnostic identification procedure. *Journal of Monetary Economics*, 52(2):381–419, March 2005a.

I. Verdinelli and L. A. Wasserman. Computing bayes factors by using a generalization of the savage-dickey density ratio. *Journal of the American Statistical Association*, 90:614–618, 1995.

V. Wieland. Monetary policy, parameter uncertainty and optimal learning. *Journal of Monetary Economics*, 46(1):199–228, August 2000.

M. Woodford. *Interest and Prices: Foundations of a Theory of Monetary Policy.* Princeton University Press, 2003a.

M. Woodford. Optimal interest rate smoothing. *Review of Economic Studies*, 70: 861–886, 2003b.

S.-C. S. Yang. A chronology of postwar u.s. federal income tax policy. Caepr Working Papers 2007-021, Center for Applied Economics and Policy Research, Economics Department, Indiana University Bloomington, Oct. 2007.

Selbständigkeitserklärung

Hiermit erkläre ich, die vorliegende Arbeit selbständig ohne fremde Hilfe verfasst und nur die angegebene Literatur und Hilfsmittel verwendet zu haben.

Alexander Kriwoluzky
1. Juni 2008